GenCore version 5.1.5

OM nucleic - nucleic search, using sw model

Run on: May 5, 2003, 18:05:53 ; Search time 1931 Seconds
(without alignments)
6118.974 Million cell updates/sec

Title: US-09-829-481-3
Perfect score: 406
Sequence: 1 ctctactacaatcactaagt.....aaaaaaaaaaaaaaaaaaa 406

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	58	14.3	4900	2	AC115591	AC115591 Dictyoste
	2	56.4	13.9	1271	9	AF090934	AF090934 Homo sapi
	3	56.2	13.8	2352	8	ATH304841	AJ304841 Arabidops
	4	56.2	13.8	156533	2	AC117070	AC117070 Dictyoste
c	5	56	13.8	204557	10	AL669851	AL669851 Mouse DNA
c	6	56	13.8	206345	2	AL844864	AL844864 Mus muscu
	7	55.8	13.7	797	10	BC010766	BC010766 Mus muscu
	8	55.8	13.7	3543	9	HSM803534	AL832227 Homo sapi
c	9	55.8	13.7	17294	6	AX345888	AX345888 Sequence
	10	55.6	13.7	1652	3	AF151644	AF151644 Caenorhab
	11	55.6	13.7	1717	3	AY121632	AY121632 Drosophil
	12	55.4	13.6	3024	9	AF146568	AF146568 Homo sapi
	13	55.4	13.6	27785	2	AC116978	AC116978 Dictyoste
	14	55	13.5	1432	3	AY069596	AY069596 Drosophil
	15	55	13.5	1805	3	DDIDPP5A	L05617 Dictyosteli
	16	55	13.5	3152	9	BC034961	BC034961 Homo sapi
	17	55	13.5	4003	9	HSM804273	AL832962 Homo sapi
c	18	55	13.5	15903	2	AC115612	AC115612 Dictyoste
	19	55	13.5	160759	2	AC117082	AC117082 Dictyoste
	20	54.8	13.5	1356	10	BC034886	BC034886 Mus muscu
c	21	54.8	13.5	93491	2	AC116967	AC116967 Dictyoste
	22	54.8	13.5	209800	2	AC121580	AC121580 Mus muscu
c	23	54.6	13.4	425	6	AX284646	AX284646 Sequence
	24	54.6	13.4	1451	17	AF118083	Af118083 Homo sapi
	25	54.6	13.4	2459	3	AY113249	AY113249 Drosophil
c	26	54.6	13.4	113880	3	PFMAL3P4	AL008970 Plasmodiu
c	27	54.6	13.4	160759	2	AC117082	AC117082 Dictyoste
	28	54.6	13.4	171187	2	AC116960	AC116960 Dictyoste
	29	54.4	13.4	1118	9	BC012595	BC012595 Homo sapi
	30	54.4	13.4	1577	9	BC011971	BC011971 Homo sapi
	31	54.4	13.4	1685	9	AK000137	AK000137 Homo sapi
c	32	54.4	13.4	2097	3	AB079606	AB079606 Dictyoste
	33	54.4	13.4	2789	10	BC026672	BC026672 Mus muscu
	34	54.4	13.4	3440	9	BC013609	BC013609 Homo sapi
	35	54.4	13.4	7028	9	HSM803514	AL832207 Homo sapi
	36	54.2	13.3	785	8	AF257780	AF257780 Populus t
	37	54.2	13.3	1327	6	AX306545	AX306545 Sequence
	38	54.2	13.3	1641	9	BC010739	BC010739 Homo sapi
	39	54	13.3	231807	2	AC131100	AC131100 Mus muscu
	40	54	13.3	283086	2	AC124723	AC124723 Mus muscu
	41	53.8	13.3	1588	17	AF116610	Af116610 Homo sapi
	42	53.8	13.3	1661	9	BC025753	BC025753 Homo sapi
	43	53.8	13.3	2005	9	HSM803426	AL832119 Homo sapi
c	44	53.8	13.3	6290	6	AX345949	AX345949 Sequence
c	45	53.8	13.3	20622	2	AC115609	AC115609 Dictyoste

ALIGNMENTS

RESULT 1
AC115591/c
LOCUS AC115591 4900 bp DNA linear HTG 21-MAR-2002
DEFINITION Dictyostelium discoideum chromosome 2 map 4713821-4718719 strain AX4, *** SEQUENCING IN PROGRESS ***, in ordered pieces.
ACCESSION AC115591
VERSION AC115591.1 GI:19569972
KEYWORDS HTG; HTGS_PHASE2.
SOURCE Dictyostelium discoideum.
ORGANISM Dictyostelium discoideum
Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE 1 (bases 1 to 4900)
AUTHORS Gloeckner,G., Eichinger,L., Szafranski,K., Pachebat,J., Dear,P., Lehmann,R., Baumgart,C., Parra,G., April,J.F., Guigo,R., Kumpf,K., Tunggal,B., Cox,E., Quail,M.A., Platzer,M., Rosenthal,A. and

```
            Noegel,A.A.
  TITLE     Sequence and Analysis of Chromosome 2 of Dictyostelium
  JOURNAL   Unpublished
  REMARK    The Dictyostelium Genome Sequencing Consortium
REFERENCE   2  (bases 1 to 4900)
  AUTHORS   Baumgart,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-MAR-2002) Genome Analysis, Institute of Molecular
            Biotechnology, Beutenberstr. 11, Jena 07745, Germany
COMMENT     CDS predictions from GeneID may contain errors. Further Information
            is available from IMB Jena, Department of Genome Analysis
            (http://genome.imb-jena.de/dictyostelium/)
            and the Univerity Colonge, Institute for Biochemistry I
            (http://www.uni-koeln.de/dictyostelium/project.shtml
            Funding
            Agency : Deutsche Forschungsgemeinschaft (DFG).
            * NOTE: This is a 'working draft' sequence.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
FEATURES             Location/Qualifiers
     source          1. .4900
                     /organism="Dictyostelium discoideum"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /chromosome="2"
                     /map="4713821-4718719"
     CDS             join(984. .1012,1234. .1359)
                     /note="ORF_ID:dd_01041"
                     /pseudo
                     /codon_start=1
     CDS             2628. .3113
                     /note="ORF_ID:dd_01042"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="AAL92276.1"
                     /db_xref="GI:19569973"
                     /translation="MTVINKELFINKKVYQDNKIENQTYECVIDHIGDEKYVILELTP
                     YVTTDSEGKPLKYLEKRYVLASTDTEAYHSGQKQWYVENTKELHKKKWRSPECLGGGI
                     MTIDQNNKSIKTFGTSYGYGDPPIEILKSILNTVYPNYNLDITITSEVRELVKPKHNG
                     Y"
     CDS             complement(join(3194. .4219,4336. .4764))
                     /note="ORF_ID:dd_01043"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="AAL92277.1"
                     /db_xref="GI:19569974"
                     /translation="METILYDTQTFCPICTIINRKGLILKPAIVVSRDDDKVYLKCCC
                     EQNHSKQFETLYCSSLKFFKRILSFDPVIGNRFNNSSIKDIESMGKYNDGNNQIKVSK
                     NPTISNFVLEFSIFQNSTFLSDDEIKFNIQQFQTFYPKNKKFVLKVIARGSNNINGIN
                     DKVKYISSLLKGYPILLEVTLERLNEIGKLSNSCLLGTKVYPAVKYYLKQGDELIFIK
                     EINQLLNVLREYSDMQCVVTVALERKFANLTDVLKLLRKSNDVIRFIILSLERPPKQL
                     VETLTIFNNNINNNNNNNNNEENENLICNINDPFELIKEIEKSTDGSINGDDFYPLNV
                     GQALEPMLNMMGYGLYSIKPSSYCSFATCLVNTQSIESRPINRLLNIGKLYYDLLPLL
                     PELEEKIGFFNGLKLKSILKSCNIPGVQLPNLFDYLTDRSKSEITRKIIDQTQILIIH
                     NNMDLASLDLKRRCNCSVETKIKDGFISSCTGCI"
BASE COUNT      2108 a    537 c    425 g   1830 t
ORIGIN

  Query Match             14.3%;  Score 58;  DB 2;  Length 4900;
  Best Local Similarity   59.9%;  Pred. No. 0.23;
  Matches   97;  Conservative    0;  Mismatches   65;  Indels    0;  Gaps    0;

Qy     242 TTTGCCATAATGGAGACCCGTTTTTATTGAATATCGTCAGTTTCCAATTAAAGTCATTTC 301
           |||   ||  ||   | |  ||||  ||  | ||  |   ||   |||||  |||||||
Db    2684 TTTATTATCTTGATAAACTTTTTTATTTATAAATAATTCTTTATTAATTACTGTCATTTT 2625

Qy     302 GAGCCATACTGAATAATTTTGTAATCTAACAACAGATGCAATAGTTTAAATAAACTTATA 361
             ||  |   |   ||| ||| | |   | |   || | ||||| |||| ||| |
Db    2624 ATTTTATTTTTTTTTTATTTTAACTTGAATTAAAAATAAAAATTTTAAGTAAAATTAAA 2565

Qy     362 CTTAACTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 403
            ||||  || |||| ||||||||||||||||||||   |||||

Db    2564 ATTAAAATTAAAAATAAAAAAAAAAAAAAAAAAAATTGAAAAA 2523


RESULT 2
AF090934
LOCUS       AF090934               1271 bp    mRNA    linear   PRI 12-JAN-2000
DEFINITION  Homo sapiens clone HQ0518 PRO0518 mRNA, complete cds.
ACCESSION   AF090934
VERSION     AF090934.1  GI:6690233
KEYWORDS    FLI_CDNA.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1271)
  AUTHORS   Yu,Y., Zhang,C., Luo,L., Ouyang,S., Zhang,S., Li,W., Wu,J.,
            Zhou,S., Liu,M. and He,F.
  TITLE     Functional prediction of the coding sequences of 50 new genes
            deduced by analysis of cDNA clones from human fetal liver
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1271)
  AUTHORS   Yu,Y., Zhang,C., Luo,L., Ouyang,S., Zhang,S., Li,W., Wu,J.,
            Zhou,S., Liu,M. and He,F.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-SEP-1998) Dept. of Experimental Hematology, Beijing
            Institute of Radiation Medcine, 27 Taiping RD, Beijing 100850,
            P.R.China
FEATURES             Location/Qualifiers
     source          1. .1271
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="HQ0518"
                     /tissue_type="liver"
                     /dev_stage="fetus"
     CDS             233. .439
                     /codon_start=1
                     /product="PRO0518"
                     /protein_id="AAF24048.1"
                     /db_xref="GI:6690234"
                     /translation="MQSRDTVNVKFTPRSFWGSRPKHTGGRSEMRAHVPQAPWARQAP
                     PVLPLWTVVNDHPHEKPVSRPQNT"
BASE COUNT      477 a    266 c    230 g    298 t
ORIGIN

  Query Match             13.9%;  Score 56.4;  DB 9;  Length 1271;
  Best Local Similarity   61.6%;  Pred. No. 0.53;
  Matches   90;  Conservative    0;  Mismatches   56;  Indels    0;  Gaps    0;

Qy     261 GTTTTTATTGAATATCGTCAGTTTCCAATTAAAGTCATTTCGAGCCATACTGAATAATTT 320
           || || | || |  || |   |||| ||| | |    |||| ||  |    || |  ||
Db     990 GTATTAAGTGCACTTCTTGTATTTCTAATAAGATGACTTTCCAGAAAGTGAGATTTGTTA 1049

Qy     321 TGTAATCTAACAACAGATGCAATAGTTTAAATAAACTTATACTTAACTTTTAAAAAAAAA 380
           |||  |        | | | || |  | ||  ||  | |||  | | | | |||||||||
Db    1050 TGTTCTGGCTTTTAAAAGGTAAAATATAAATAAATTTCATAACTTAATCTAAAAAAAAAA 1109

Qy     381 AAAAAAAAAAAAAAAAAAAAAAAAAA 406
           ||||||||||||||||||||||||||
Db    1110 AAAAAAAAAAAAAAAAAAAAAAAAAA 1135


RESULT 3
ATH304841
LOCUS       ATH304841              2352 bp    mRNA    linear   PLN 23-DEC-2000
DEFINITION  Arabidopsis thaliana mRNA for dynamin-like protein ADLP1 (dlp1
            gene).
ACCESSION   AJ304841
VERSION     AJ304841.1  GI:11991505
KEYWORDS    dlp1 gene; dynamin-like protein DLP1.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
```

```
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 2352)
  AUTHORS   Jasper,F. and Menzel,D.
  TITLE     Identification of a subgroup of closely related dynamin-like
            proteins in Arabidopsis thaliana
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2352)
  AUTHORS   Jasper,F.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-DEC-2000) Jasper F., Zellbiologie der Pflanzen,
            Botanisches Institut, Universitaet Bonn, Kirschallee 1, Bonn,
            53115, GERMANY
FEATURES             Location/Qualifiers
     source          1. .2352
                     /organism="Arabidopsis thaliana"
                     /cultivar="Landsberg erecta"
                     /db_xref="taxon:3702"
                     /tissue_type="root"
     gene            201. .2045
                     /gene="dlp1"
     CDS             201. .2045
                     /gene="dlp1"
                     /codon_start=1
                     /product="dynamin-like protein DLP1"
                     /protein_id="CAC19656.1"
                     /db_xref="GI:11991506"
                     /translation="MATMKSLIGLINKIQRACTVLGDHGGEGMSLWEALPTVAVVGGQ
                     SSGKSSVLESVVGRDFLPRGSGIVTRRPLVLQLHKTEDGTTEYAEFLHAPKKRFADFA
                     AVRKEIEDETDRITGKSKQISNIPIQLSIYSPNVVNLTLIDLPGLTKVAVDGQPESIV
                     QDIENMVRSYVEKPNCIILAISPANQDIATSDAIKLAREVDPTGERTFGVATKLDIMD
                     KGTDCLDVLEGRSYRLQHPWVGIVNRSQADINKRVDMIAARRKEQEYFETSPEYGHLA
                     SRMGSEYLAKLLSQHLETVIRQKIPSIVALINKSIDEINAELDRIGRPIAVDSGAQLY
                     TILELCRAFDRVFKEHLDGGRPGGDRIYGVFDHQLPAALKKLPFDRHLSTKNVQKVVS
                     EADGYQPHLIAPEQGYRRLIDGSISYFKGPAEATVDAVHFVLKELVRKSISETEELKR
                     FPTLASDIAAAANEALERFRDESRKTVLRLVDMESSYLTVEFFRKLHLEPEKEKPNPR
                     NAPAPNADPYSDNHFRKIGSNVSAYINMVCDTLRNSLPKAVVYCQVREAKRSLLNFFY
                     AQVGRKEKEKLGAMLDEDPQLMERRGTLAKRLELYKQARDDIDAVAWK"
BASE COUNT      706 a    472 c    529 g    645 t
ORIGIN

  Query Match             13.8%;  Score 56.2;  DB 8;  Length 2352;
  Best Local Similarity   58.8%;  Pred. No. 0.54;
  Matches   97;  Conservative    0;  Mismatches   68;  Indels    0;  Gaps    0;

Qy     237 ATTTATTTGCCATAATGGAGACCCGTTTTTATTGAATATCGTCAGTTTCCAATTAAAGTC 296
           |||| | |  ||    |    ||   |||||  ||  |||| ||  ||    | ||  |
Db    2188 ATTTCTATCACAACTTTCCCACTTTTTTTTGGTGTGTATCTTCGATTCTTGAATACTGGT 2247

Qy     297 ATTTCGAGCCATACTGAATAATTTTGTAATCTAACAACAGATGCAATAGTTTAAATAAAC 356
            ||| || |  |   ||   ||||| | ||  ||     |   | |    ||  || |
Db    2248 TTTTAGAACAGTTAAGACATATTTTGAAGTCGCACTCTTTACATTAGACAAAAAGGAAGC 2307

Qy     357 TTATACTTAACTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 401
           | ||  || | |||||  ||||||||||||||||||||||||||||
Db    2308 TAATCTTTGATTTTTAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2352


RESULT 4
AC117070
LOCUS       AC117070               156533 bp    DNA     linear   HTG 06-APR-2002
DEFINITION  Dictyostelium discoideum chromosome 2 map 2268371-2355204 strain
            AX4, *** SEQUENCING IN PROGRESS ***, in ordered pieces.
ACCESSION   AC117070
VERSION     AC117070.1  GI:20066187
KEYWORDS    HTG; HTGS_PHASE2.
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 156533)
  AUTHORS   Gloeckner,G., Eichinger,L., Szafranski,K., Pachebat,J., Dear,P.,
            Lehmann,R., Baumgart,C., Parra,G., April,J.F., Guigo,R., Kumpf,K.,
            Tunggal,B., Cox,E., Quail,M.A., Platzer,M., Rosenthal,A. and
            Noegel,A.A.
  TITLE     Sequence and Analysis of Chromosome 2 of Dictyostelium
  JOURNAL   Unpublished
  REMARK    The Dictyostelium Genome Sequencing Consortium
REFERENCE   2  (bases 1 to 156533)
  AUTHORS   Baumgart,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-APR-2002) Genome Analysis, Institute of Molecular
            Biotechnology, Beutenberstr. 11, Jena 07745, Germany
COMMENT     CDS predictions from GeneID may contain errors. Further Information
            is available from IMB Jena, Department of Genome Analysis
            (http://genome.imb-jena.de/dictyostelium/)
            and the Univerity Colonge, Institute for Biochemistry I
            (http://www.uni-koeln.de/dictyostelium/project.shtml
            Funding
            Agency : Deutsche Forschungsgemeinschaft (DFG).
            * NOTE: This is a 'working draft' sequence.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
FEATURES             Location/Qualifiers
     source          1. .156533
                     /organism="Dictyostelium discoideum"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /chromosome="2"
                     /map="2268371-2355204"
     CDS             join(3827. .3945,5512. .5557)
                     /note="ORF_ID:dd_00443"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="AAM09328.1"
                     /db_xref="GI:20066214"
                     /translation="MSADQPPLRTMLIMSADQPPMQISQIMSVDPTTLITIHSVGYLG
                     ASLDEGNQTA"
     CDS             complement(7554. .8315)
                     /note="ORF_ID:dd_00439"
                     /codon_start=1
                     /product="Discoidin I, D chain"
                     /protein_id="AAM09325.1"
                     /db_xref="GI:20066211"
                     /translation="MSTQGLVTLLGNAQCHLRTSTNYNGVHTQFNAALNYKNKGTNTI
                     DGSEAWCSSIVDTNQYIVAGCEVPRTFMCVALQGRGDHDQWVTSYKIRYSLDNVTWSE
                     YRNGAAITGVTDRNTVVNHFFDTPIRARSIAIHPLTWNNHISLRCEFYTQPVQSSVTQ
                     VGADIYTGDNCALNTGSGKREVVVPVKFQFEFATLPKVALNFDQIDCTDATNQTRIGV
                     QPRNITTKGFDCVFYTWNENKVYSLRADYIATALE"
     CDS             complement(9039. .9800)
                     /note="ORF_ID:dd_00438"
                     /codon_start=1
                     /product="Discoidin I, C chain and B chain."
                     /protein_id="AAM09326.1"
                     /db_xref="GI:20066212"
                     /translation="MSTQGLVQLISNAQCHLRTSTNYNDVHTQFNAVLNYKNKGTNTI
                     DGSEAWCSSIVDTNQYIVAGCEVPRTFMCVALQGRGDHDQWVTSYKIRYSLDNVTWSE
                     YRNGAAITGVTDRNTVVNHFFDTPIRARSIAIHPLTWNNHISLRCEFYTQPVQSSVTQ
                     VGADIYTGDNCALNTGSGKREVVVPVKFQFEFATLPKVALNFDQIDCTDATNQTRIGV
                     QPRNITTKGFDCVFYTWNANKVYSLRADYIATALE"
     CDS             join(12482. .12692,13544. .13672,14278. .14315)
                     /note="ORF_ID:dd_00437"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="AAM09327.1"
                     /db_xref="GI:20066213"
                     /translation="MIKEEDEQLIKKEVKSISIEDDLYDEDSYCLDCGQFKTHYSRHC
                     SSIHQQVNVNKVIMDHCKRYPRESFRARDAEWDRTRKIVEKQLRQELRCTDNRESINR
                     TASIKRRLEKFKYVASSIPILKN"
     CDS             14598. .15479
                     /note="ORF_ID:dd_00448"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="AAM09329.1"
                     /db_xref="GI:20066215"
                     /translation="MNFFDKLKGELNLGEEELNKKTIQIRLFYPHTLELNSNSTKLID
                     NFIVNGTLNIVRRQKGITVFGRILKEKFTLLHECYSTFKYEDFPLQKFILYTPGYHYV
```

```
                HFYTFTKNIIKKINNQMVKENWGLEPLLIDECNIVMLPDVIRYSCILLFNKEDDISII
                KTNIDAIKGSKHPIRAFISEYNNKEIRDVRKKKNTSNKNQQPTTTNTTKPIPSTPPTT
                NASTTTPITTTPITTTQTTTTPTTITHQLPTSNTTINNLNKTTTTTTTTTTTTKPINK
                ITTSNGTKFPIASSIRK"
CDS             15629. .18214
                /note="ORF_ID:dd_00447"
                /codon_start=1
                /product="Putative Retron-type reverse transcriptase"
                /protein_id="AAM09330.1"
                /db_xref="GI:20066216"
                /translation="MFNFERIDHGSGKGTGIAIENRDTRKGHISINFKDDDGRILSIK
                YNSFNSINILLINAPATISERNTFIINSKSLFKKYNSINHQIIAGDFNNNHDCNSFFG
                TELRKIIDQDMLLDTGIEENTPTFPRSMKRLDRIYCHPTLLNQNSKLVVHNTVFNKSD
                HFPITITIQTNRETTTTTTTTKLERLPWTLCKEILNNKHIHDGLSELISKNKDKIKSV
                EEWTKFKNNVIRDYLKKEQNKIKKEKNKRKYVIHKLLGNSDIIPKMRKELNEEISRIL
                EEERKVKAWDIKLKLHLHQETPSKYLTSILKSRAKDKSIFQIKDKDNKTISDKENIAK
                RFVEFYQDQYEEKEDNEETHKKLLEKWEVDVDLIKKLEIDRPIRIHEVTKAIKTSSIH
                KSPGLDGINALFYKYHINSIARILTIAFNDLLTNKKEIPTKFKEGVITTIFKKGDELN
                ISNRRPITLLNTDYKILSKILNSRLLDITSKIINKFQNGFVPNRFIQDNIQIMKEVIE
                ISNKRKNNTLITFYDFNKAFDSISHKSITRTLEHIGIPPKFTAILLNLLKDTKNKIKI
                NDFLVNGITIRTGTKQGDPISPTIFALVLEPLLIDIINDNTIKGFTLPNSKSLKLTAF
                ADDIATFTNSTEELMKINTKIQKYCSATSSSLNKEKTVMIAIGDKPHDLPFQESTVPE
                RYLGLNFTKTGLNSKYNTLIQEMKNNLIKWKSQAITMKAKMTILKTYVLSKLTYHQYM
                DNLNEEQIEEINNMTRWFLFSSVKNTYTEERKYKTMMKIDRAYADWKEGGIKLWDIEL
                RHIAFKIWYMNRLLHNNYNNNNNTLQEWYMEQLSRKKPTLQPSTICADTGVYSESNFT
                KTIQR"
CDS             join(18321. .18339,18404. .18924)
                /note="ORF_ID:dd_00446"
                /codon_start=1
                /product="hypothetical protein"
                /protein_id="AAM09331.1"
                /db_xref="GI:20066217"
                /translation="MAEVMGGRNTLFRFFSRSLPGINHERDTRCKICGHLFRDPYSHL
                FTLCQDILDIEKTIISTVNKLSFIKIQRWSMDTLDISKYNRTERIFPNLIGIIAHQLW
                KIICHKLFNTDESKPEPKFEQKVIETELLNLIETEKFITLKKIKHDEAILKNTNQDLH
                KYKFNKAWQTPAAPNPLPI"
CDS             19886. .20788
                /note="ORF_ID:dd_00444"
                /codon_start=1
                /product="Putative Nuclear protein"
                /protein_id="AAM09332.1"
                /db_xref="GI:20066218"
                /translation="MNNIFLIISSIILSIVIFVSGECAIDFSSEISVGISDSQWSCLA
                SNNQRVIIQVWSGGGQYNSNISSVVSAAEQAGFDNIDLYAFLCSECDGNYPASSAIQS
                LVSSLKSDGINFNMLWIDVEQCDGCWGAESDNADYVQEAVETAQGLGVLVGVYSSEGE
                WPQTVGNLSTLSQYPLWYAHYDDNPSFSDTAFYEFGGWTSPAMKQYIGNTNQCGVSVD
                LDFYGSGSGCSTSSGSASGSASGSASGSASGSNSGSSNSGSSNSGSSNSGSNSGSSNS
                GSGNSGSSNSGSASGSGTGSGSSI"
CDS             complement(join(21433. .23231,23387. .23705))
                /note="ORF_ID:dd_00442"
                /codon_start=1
                /product="hypothetical protein"
                /protein_id="AAM09333.1"
                /db_xref="GI:20066219"
                /translation="MSINKEKQKIINEINNEPKDNISSGDNNIIKIRETEGNNEDCTD
                IPTPTPIPIIDVELISPLPSPPPPPTTIDNNNKTTPIITPQLIDFDLPSSATILHSPF
                TNSTIILIGSVHIHKGSSDEVSEIIRKWKPDTVFVELCSSRAGIIFNTNNDDSDYYYN
                NNNNNNKNNNTSNKSKRTKKILDGRTKSVYSRRLPQREKQHFLSDLDDDQFSSASTSK
                FVLETEIQDIDIEDMTQEDNELYTSEETEQDENFDDDDDYYGFDYLDHNDNNINHLMN
                QPRNFNFYLYDDDNDLIEENNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNK
                NKNSDNLKTTKLINNNFKNNIFNSDNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
                FINNINLNNSEFNYSSDADDERSHSSFESITTFTTQTYIDNEEEEEEDNEEENEENNE
                EIKNDSQESISIKDIIRIIKCNGLSGVLHILIAKLINKAGKKSKVGPGSEFITAFLEA
                RKIGSLVVLGDRQVGITLQRVWNSLSWLEKIKFVFYLFMASLSEITTEEIDAIKNSSD
                ELVNKLLNEFRGRFPSVVHTIVTERDQYMAARLRMCPGKKIVAVVGKGHIGGIIREWE
                NNNINLFELESVNIIKKSNLNNNNNNNSNNNFNKQNQQQQQQQQQQQSSKSSWSTVTK
                WASLILIPIGIGLYFFNKKFNFI"
CDS             complement(25734. .26420)
                /note="ORF_ID:dd_01644"
                /codon_start=1
                /product="hypothetical protein"
                /protein_id="AAM09335.1"
                /db_xref="GI:20066221"
                /translation="MFEINKIYMDNASEDERKIISKSKIQKIDKEILPVCASPASNVN
                DIDNFQNKLNNQLNEKKKKIKNQGIVSIDSNSLPIKNKNELIQFNKLLSESNGNLKNE
                KTESFINDKLIVDKDKNLYGVGRYVDHVYCVFGKYQLKKSVIKVTKPDPNIPISSLDL
                LAIYTKAYQYMYKEQKQWFDYKKTRINISNSPHGIYRYGINDIHYNGLSTIIEFDTFL
                VCSFDVDPIK"
CDS             complement(27154. .27738)
                /note="ORF_ID:dd_01643"
                /codon_start=1
                /product="hypothetical protein"
                /protein_id="AAM09336.1"
                /db_xref="GI:20066222"
                /translation="MENYLFRDANSEELKEIEGLKPIKKLENILKDIGAPNKQITTLE
                NYKNYLKEDALSKKSWGLESFNQPNQTDDDLFKETVLIEGGDNTKHVYCFIDAYTANP
                VTIKVTRPDANSLQPLTNVELLAIYSKAFQFIYEEEEKEVGNPGHIQGMLNRGVSHGR
                YGIWGHDLSDLVYNGFSDVLIYKDFIFCEFRVDS"
CDS             complement(join(28103. .28128,28441. .28881,28964. .29004,
                29963. .30505,30619. .30989))
                /note="ORF_ID:dd_01642"
                /codon_start=1
                /product="Putative GTPase"
                /protein_id="AAM09337.1"
                /db_xref="GI:20066223"
                /translation="MKYINLIFILICFYINYQSIGADRNEFPINEILAAKSILKSLYN
                FEAKTLEEICFYGCFVCSPIYEPKTYGITTIRINSTRLSFVSYDFSIFRYLSELIIEE
                NIQIQTVFYEKTLPLLKNLELLEVSKQEQPFPDNFSPPEKLKTVQFNSISVPLSSIWF
                EGVINRIIVKKTLPGFKYPKLTKINTYINTLSLSLSNHIDTNVPSMDLFPNLELINFQI
                HNEMSQKGYKNFSIDSINQDWYKRVNAMEIQFINSGKEATIQKFTLQQSLISRFKLDY
                LLIDGIGFTVDPSIGYLNFSKMTDKGFVLHLFKFVLAGDSGVGKTSILFRITDDTFTE
                THITIGIEFKIKTVYIEGKPIKLQIWDTAGEKRFRVHNSHYRGCHGVIIVYDVTDQRS
                FENVPSWIEDIRRYANENVIKIIIGNKNDLVSQKVVDPFLAQEFADSLDIPFKEISAK
                QSINIEEAFISLVKLFLKTNNNT"
CDS             complement(32023. .32289)
                /note="ORF_ID:dd_01641"
                /codon_start=1
                /product="hypothetical protein"
                /protein_id="AAM09338.1"
                /db_xref="GI:20066224"
                /translation="MKCLKKVTRIFQLIQLIEYINVALEFNFINSGNEATIQKFTSQQ
                SLISKFKLCWLVFKGISFTLDPSIGYLFNNFSKMDDIYGIDLYV"
CDS             complement(join(34992. .35316,35500. .35717))

  Query Match          13.8%;  Score 56.2;  DB 2;  Length 156533;
  Best Local Similarity  54.0%;  Pred. No. 0.32;
  Matches  115;  Conservative    0;  Mismatches   98;  Indels    0;  Gaps    0;

Qy     194 GGAACTTTCAGGACAACCTGCGTTTGCTATAGGTGAAAATCCGATTTATTTGCCATAATG 253
           |||  | | |   ||      || |  || | |  ||     |  |||||||| ||
Db   32418 GGATATATAAAACCAGGTAATGTATTATAAACATATAATGAATAAATATTTGCCTAAAAC 32477

Qy     254 GAGACCCGTTTTTATTGAATATCGTCAGTTTCCAATTAAAGTCATTTCGAGCCATACTGA 313
           | |     || || | | |     |  |  |||| || || |  | |   ||   |
Db   32478 CAAATTGAACTTAATGGTACAAAAATTGAATTAAATTGAACTCCTGACAATTTATCTGGG 32537

Qy     314 ATAATTTTGTAATCTAACAACAGATGCAATAGTTTAAATAAACTTATACTTAACTTTTAA 373
           ||| |    | ||||   ||  ||||  | | ||  | |    | || | |  |  ||
Db   32538 ATATTAAATTTATCTGGGAAAGGATGTTCTTGGTTTGAAATTTCTCTATTAATTATAAAA 32597

Qy     374 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 406
           ||||||||||||||||||||||||||||||||
Db   32598 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 32630


RESULT 5
AL669851/c
LOCUS       AL669851              204557 bp    DNA     linear   ROD 21-JUN-2002
DEFINITION  Mouse DNA sequence from clone RP23-18K13 on chromosome X, complete
            sequence.
ACCESSION   AL669851
VERSION     AL669851.8  GI:21540074
KEYWORDS    HTG.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 204557)
  AUTHORS   Howden,P.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-JUN-2002) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Jun 23, 2002 this sequence version replaced gi:20502249.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence was finished as follows unless otherwise noted: all
            regions were either double-stranded or sequenced with an alternate
            chemistry or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by at least
            one plasmid subclone or more than one M13 subclone; and the
            assembly was confirmed by restriction digest. The following
            abbreviations are used to associate primary accession numbers given
            in the feature table with their source databases: Em:, EMBL; Sw:,
            SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information on the WORMPEP
            database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-18K13 is
            from the RPCI-23 Mouse PAC Library
            constructed by the group of Pieter de Jong.
            For further details see http://www.chori.org/bacpac/home.htm
            VECTOR: pBACe3.6
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            --------------.
FEATURES             Location/Qualifiers
     source          1. .204557
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /chromosome="X"
                     /clone="RP23-18K13"
                     /clone_lib="RPCI-23"
BASE COUNT    62833 a  37986 c  37959 g  65779 t
ORIGIN

  Query Match             13.8%;  Score 56;  DB 10;  Length 204557;
  Best Local Similarity   55.7%;  Pred. No. 0.34;
  Matches  107;  Conservative    0;  Mismatches   85;  Indels    0;  Gaps    0;

Qy     215 GTTTGCTATAGGTGAAAATCCGATTTATTTGCCATAATGGAGACCCGTTTTTATTGAATA 274
           ||||||| |    | | ||       |||||  ||  |||  | |  | | |     |||
Db  161602 GTTTGCTCTTCAAGGACATTGAGAAGATTTGTTATCTTGGCCAGCTCTCTCTGGAATATA 161543

Qy     275 TCGTCAGTTTCCAATTAAAGTCATTTCGAGCCATACTGAATAATTTTGTAATCTAACAAC 334
           | ||   | ||| ||| |||||| |  | |||    ||| | | |   |    |
Db  161542 GCATCCTGTACCATTTAGAGTCATGTGAAACCAAGGGGAAAACTATGTGACAGAACATCT 161483

Qy     335 AGATGCAATAGTTTAAATAAACTTATACTTAACTTTTAAAAAAAAAAAAAAAAAAAAAAA 394
           ||   |  ||    |||||| |    || |  | |  || | | ||||||||||||||||
Db  161482 TGAAATACAAGAGGCAATAAAGTCTGCCTAAGATCTAGAACAGATAAAAAAAAAAAAAAA 161423

Qy     395 AAAAAAAAAAAA 406
           ||||||||||||
Db  161422 AAAAAAAAAAAA 161411


RESULT 6
AL844864/c
LOCUS       AL844864              206345 bp    DNA     linear   HTG 29-JUL-2002
DEFINITION  Mus musculus chromosome X clone RP23-360L8, *** SEQUENCING IN
            PROGRESS ***, 15 unordered pieces.
ACCESSION   AL844864
VERSION     AL844864.2  GI:22022618
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Mus musculus.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 206345)
  AUTHORS   McLay,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-JUL-2002) Wellcome Trust Sanger Institute, Hinxton,
            Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
            humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
COMMENT     On Jul 30, 2002 this sequence version replaced gi:22003281.
            -------------- Genome Center
            Center: Wellcome Trust Sanger Institute
            Center code: SC
            Web site: http://www.sanger.ac.uk
            Contact: humquery@sanger.ac.uk
            -------------- Project Information
            Center project name: bM360L8
            -------------- Summary Statistics
            Assembly program: XGAP4; version 4.5
            Chemistry: Dye-terminator; 100% of reads
            Consensus quality: 202112 bases at least Q40
            Consensus quality: 203382 bases at least Q30
            Consensus quality: 204223 bases at least Q20
            Insert size: 204945; sum-of-contigs
            Insert size: 202604; 0.4% error; agarose-fp
            Quality coverage: 6.25x in Q20 bases; sum-of-contigs Quality
            coverage: 6.39x in Q20 bases; agarose-fp
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 15 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     5919: contig of 5919 bp in length
            *     5920  6019: gap of      100 bp
            *     6020    45539: contig of 39520 bp in length
            *    45540 45639: gap of      100 bp
            *    45640    49244: contig of 3605 bp in length
            *    49245 49344: gap of      100 bp
            *    49345    52863: contig of 3519 bp in length
            *    52864 52963: gap of      100 bp
            *    52964    56198: contig of 3235 bp in length
            *    56199 56298: gap of      100 bp
            *    56299    64095: contig of 7797 bp in length
            *    64096 64195: gap of      100 bp
            *    64196    76544: contig of 12349 bp in length
            *    76545 76644: gap of      100 bp
            *    76645   104869: contig of 28225 bp in length
            *   104870 104969: gap of      100 bp
            *   104970   113702: contig of 8733 bp in length
            *   113703 113802: gap of      100 bp
            *   113803   157886: contig of 44084 bp in length
            *   157887 157986: gap of      100 bp
            *   157987   174948: contig of 16962 bp in length
            *   174949 175048: gap of      100 bp
            *   175049   185910: contig of 10862 bp in length
            *   185911 186010: gap of      100 bp
            *   186011   190330: contig of 4320 bp in length
            *   190331 190430: gap of      100 bp
            *   190431   198187: contig of 7757 bp in length
            *   198188 198287: gap of      100 bp
            *   198288   206345: contig of 8058 bp in length.
FEATURES             Location/Qualifiers
     source          1. .206345
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /chromosome="X"
```

```
                     /clone="RP23-360L8"
                     /clone_lib="RPCI-23"
     misc_feature    1. .5919
                     /note="assembly_fragment:01858
                     fragment_chain:1
                     clone_end:SP6
                     vector_side:left"
     misc_feature    6020. .45539
                     /note="assembly_fragment:02230
                     fragment_chain:1"
     misc_feature    45640. .49244
                     /note="assembly_fragment:00122
                     fragment_chain:1"
     misc_feature    49345. .52863
                     /note="assembly_fragment:01053
                     fragment_chain:2"
     misc_feature    52964. .56198
                     /note="assembly_fragment:00150
                     fragment_chain:2"
     misc_feature    56299. .64095
                     /note="assembly_fragment:01584
                     fragment_chain:2"
     misc_feature    64196. .76544
                     /note="assembly_fragment:01500
                     fragment_chain:2"
     misc_feature    76645. .104869
                     /note="assembly_fragment:01351
                     fragment_chain:3"
     misc_feature    104970. .113702
                     /note="assembly_fragment:02015
                     fragment_chain:3"
     misc_feature    113803. .157886
                     /note="assembly_fragment:01814
                     fragment_chain:3"
     misc_feature    157987. .174948
                     /note="assembly_fragment:02307
                     fragment_chain:3"
     misc_feature    175049. .185910
                     /note="assembly_fragment:00195
                     fragment_chain:4"
     misc_feature    186011. .190330
                     /note="assembly_fragment:00963
                     fragment_chain:4"
     misc_feature    190431. .198187
                     /note="assembly_fragment:00683"
     misc_feature    198288. .206345
                     /note="assembly_fragment:00907"
BASE COUNT    64596 a  39135 c  39792 g  61421 t   1401 others
ORIGIN

  Query Match            13.8%;  Score 56;  DB 2;  Length 206345;
  Best Local Similarity  55.7%;  Pred. No. 0.34;
  Matches  107;  Conservative    0;  Mismatches   85;  Indels    0;  Gaps    0;

Qy     215 GTTTGCTATAGGTGAAAATCCGATTTATTTGCCATAATGGAGACCCGTTTTTATTGAATA 274
           ||||||| |    | | ||       |||||  ||  |||  | |  | | |     |||
Db   26637 GTTTGCTCTTCAAGGACATTGAGAAGATTTGTTATCTTGGCCAGCTCTCTCTGGAATATA 26578

Qy     275 TCGTCAGTTTCCAATTAAAGTCATTTCGAGCCATACTGAATAATTTTGTAATCTAACAAC 334
           | ||   | ||| ||| |||||| |  | |||    ||| | | |   |   |
Db   26577 GCATCCTGTACCATTTAGAGTCATGTGAAACCAAGGGGAAAACTATGTGACAGAACATCT 26518

Qy     335 AGATGCAATAGTTTAAATAAACTTATACTTAACTTTTAAAAAAAAAAAAAAAAAAAAAAA 394
           ||   |  ||    |||||| |    || |  | |  || | | ||||||||||||||
Db   26517 TGAAATACAAGAGGCAATAAAGTCTGCCTAAGATCTAGAACAGATAAAAAAAAAAAAAAA 26458

Qy     395 AAAAAAAAAAAA 406
           ||||||||||||
Db   26457 AAAAAAAAAAAA 26446


RESULT 7
BC010766
```

```
LOCUS       BC010766                797 bp    mRNA    linear   ROD 07-AUG-2002
DEFINITION  Mus musculus, Similar to ATP synthase, H+ transporting,
            mitochondrial F0 complex subunit F, clone MGC:18567 IMAGE:4219715,
            mRNA, complete cds.
ACCESSION   BC010766
VERSION     BC010766.1  GI:14789693
KEYWORDS    MGC.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 797)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-JUL-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey E. Green, M.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www-shgc.stanford.edu
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 25 Row: h Column: 1
            This clone was selected for full length sequencing because it
            passed the following selection criteria: matched mRNA gi: 7949004.
FEATURES             Location/Qualifiers
     source          1. .797
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /map="FVB/N"
                     /clone="MGC:18567 IMAGE:4219715"
                     /tissue_type="Colon, normal. 5 month old male mouse."
                     /clone_lib="NCI_CGAP_Co24"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     CDS             345. .671
                     /codon_start=1
                     /product="Similar to ATP synthase, H+ transporting,
                     mitochondrial F0 complex subunit F"
                     /protein_id="AAH10766.1"
                     /db_xref="GI:14789694"
                     /translation="MVLQRIFRLSSVLRSAVSVHLKRNIGVTAVAFNKELDPVQKLFV
                     DKIREYKSKRQASGGPVDIGPEYQQDLDRELYKLKQMYGKGEMDTFPTFKFDDPKFEV
                     IDKPQS"
BASE COUNT      226 a    199 c    209 g    163 t
ORIGIN

  Query Match            13.7%;  Score 55.8;  DB 10;  Length 797;
  Best Local Similarity  72.7%;  Pred. No. 0.73;
  Matches   72;  Conservative    0;  Mismatches   27;  Indels    0;  Gaps    0;

Qy     308 TACTGAATAATTTTGTAATCTAACAACAGATGCAATAGTTTAAATAAACTTATACTTAAC 367
           || |   | ||    || |   ||||  ||  || | || |||||||| |  | || ||
Db     694 TAATTTGTCATGAATTAGTTGTACAACTAATCAAAAAATTCAAATAAACATTCATTTCAC 753

Qy     368 TTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 406
           ||||||||||||||||||||||||||||||||||||||||
Db     754 AGTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 792


RESULT 8
HSM803534
```

```
LOCUS       HSM803534               3543 bp    mRNA    linear   PRI 10-JUL-2002
DEFINITION  Homo sapiens mRNA; cDNA DKFZp686P1536 (from clone DKFZp686P1536).
ACCESSION   AL832227
VERSION     AL832227.1  GI:21732773
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3543)
  AUTHORS   Wambutt,R., Heubner,D., Mewes,H.W., Weil,B., Amid,C. and Wiemann,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JUL-2002) 1, D-85764 Neuherberg, GERMANY
COMMENT     Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
            consortium of the German Genome Project.
            This clone (DKFZp686P1536) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
            information about the clone and the sequencing project is available
            at http://mips.gsf.de/proj/cDNA/.
FEATURES             Location/Qualifiers
     source          1. .3543
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="DKFZp686P1536"
                     /tissue_type="cDNA-collection"
                     /clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host
                     DH10B; sites SfiIA + SfiIB"
                     /dev_stage="adult"
     polyA_site      3394
BASE COUNT     1077 a    656 c    623 g   1187 t
ORIGIN

  Query Match             13.7%;  Score 55.8;  DB 9;  Length 3543;
  Best Local Similarity   57.0%;  Pred. No. 0.6;
  Matches  102;  Conservative    0;  Mismatches   77;  Indels    0;  Gaps    0;

Qy     228 GAAAATCCGATTTATTTGCCATAATGGAGACCCGTTTTTATTGAATATCGTCAGTTTCCA 287
           || ||  || | | | |   ||  || |||   |  |||||   |||     ||    ||
Db    3316 GATAAATCGGTGTGTATTTTATTTTGCAGAAATGGATTTATATCATACACAAAGCACACA 3375

Qy     288 ATTAAAGTCATTTCGAGCCATACTGAATAATTTTGTAATCTAACAACAGATGCAATAGTT 347
           | || |   |||||    || |   || ||       ||   || || | |   || |
Db    3376 AATATATATATTTCTTTGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3435

Qy     348 TAAATAAACTTATACTTAACTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 406
            ||| |||   | |   ||     ||||||||||||||||||||||||||||||||||
Db    3436 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3494


RESULT 9
AX345888/c
LOCUS       AX345888               17294 bp    DNA     linear   PAT 01-FEB-2002
DEFINITION  Sequence 959 from Patent WO0200928.
ACCESSION   AX345888
VERSION     AX345888.1  GI:18493774
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Olek,A., Piepenbrock,C. and Berlin,K.
  TITLE     Diagnosis of diseases associated with the immune system
  JOURNAL   Patent: WO 0200928-A 959 03-JAN-2002;
            Epigenomics AG (DE)
FEATURES             Location/Qualifiers
     source          1. .17294
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="chemically treated genomic DNA (Homo sapiens)"
BASE COUNT     4922 a    203 c   3691 g   8478 t
ORIGIN

  Query Match             13.7%;  Score 55.8;  DB 6;  Length 17294;
  Best Local Similarity   53.4%;  Pred. No. 0.5;
  Matches  117;  Conservative    0;  Mismatches  102;  Indels    0;  Gaps    0;

Qy     184 CTTTTGCAGAGGAACTTTCAGGACAACCTGCGTTTGCTATAGGTGAAAATCCGATTTATT 243
           ||||  |  |  ||     |  | || |  |  ||  | ||  | |  |||   |  |||
Db     280 CTTTAACCTATAAAAAAAAAAAATAAACCCCTATTAATTTAAATCACCATCAATTCGATT 221

Qy     244 TGCCATAATGGAGACCCGTTTTTATTGAATATCGTCAGTTTCCAATTAAAGTCATTTCGA 303
           | | || |     ||   ||||| ||   | |  |   |||   | | ||       | |
Db     220 TTCTATTAATCGTACAAATTTTTTTTTTTTTTTTTTTTTTTTTACTCAATATCAAACCA 161

Qy     304 GCCATACTGAATAATTTTGTAATCTAACAACAGATGCAATAGTTTAAATAAACTTATACT 363
            ||  | | ||| |  ||  ||   |  || | ||  |  |   |||| ||| |||
Db     160 ACCTAAATAAATTAACTTTCAAAAAATTAAAATATAAACCATAATAAAAAAAAAAATAAC 101

Qy     364 TAACTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 402
           |||  |   ||||||||||||||||||||||||||| || |||
Db     100 TAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAACAACAAA 62


RESULT 10
AF151644
LOCUS       AF151644               1652 bp    mRNA    linear   INV 27-JUL-2001
DEFINITION  Caenorhabditis elegans ionotropic GABA receptor subunit UNC-49C
            (unc-49C) mRNA, complete cds.
ACCESSION   AF151644
VERSION     AF151644.1  GI:5354173
KEYWORDS    .
SOURCE      Caenorhabditis elegans.
  ORGANISM  Caenorhabditis elegans
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
            Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE   1  (bases 1 to 1652)
  AUTHORS   Bamber,B.A., Beg,A.A., Twyman,R.E. and Jorgensen,E.M.
  TITLE     The Caenorhabditis elegans unc-49 locus encodes multiple subunits
            of a heteromultimeric GABA receptor
  JOURNAL   J. Neurosci. 19 (13), 5348-5359 (1999)
  MEDLINE   99307327
   PUBMED   10377345
REFERENCE   2  (bases 1 to 1652)
  AUTHORS   Bamber,B.A. and Jorgensen,E.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-MAY-1999) Biology, University of Utah, 257S 1400E,
            Salt Lake City, UT 84112-0840, USA
FEATURES             Location/Qualifiers
     source          1. .1652
                     /organism="Caenorhabditis elegans"
                     /strain="Bristol N2"
                     /db_xref="taxon:6239"
                     /chromosome="III"
                     /map="between lin-19 and mel-23"
     gene            1. .1652
                     /gene="unc-49C"
     CDS             114. .1460
                     /gene="unc-49C"
                     /codon_start=1
                     /product="ionotropic GABA receptor subunit UNC-49C"
                     /protein_id="AAD42386.1"
                     /db_xref="GI:5354174"
                     /translation="MARPFTLIVLLSAHLCLHVVVTQDEDSHINTQLLSSVLDRLTNR
                     TTYDKRLRPRYGEKPVDVGITIHVSSISAVSEVDMDFTLDFYMRQTWQDPRLAFGSLD
                     LGLSKEIDSLTVGVDYLDRLWKPDTFFPNEKKSFFHLATTHNSFLRIEGDGTVYTSQR
                     LTVTATCPMDLKLFPMDSQHCKLEIESYAYSTAEIEYKWCTSKEPNCSTAVKADANIE
                     LSSYKFTKICQKRTLASTSSGTYSRLRVSFIFDRDSGFYFLQIFFPASLVVVLSWISF
                     WINRDSAPSRTLIGTMTVLTETHLMTGTNRRLPPVAYVKAVDVFLGFCYLLVILALIE
                     YACVAYSKKKNEDRRRREKKTEHKPAPPTPDILHDVRLAECTCNAAPTSIIAVIKQSN
                     RFCVSHSHIDIVSRAAFPLVFILFNTLFWLILLYKSKRLPYISEHEGDRCDAPDLH"
BASE COUNT      456 a    413 c    323 g    460 t
ORIGIN
```

```
  Query Match            13.7%;  Score 55.6;  DB 3;  Length 1652;
  Best Local Similarity  55.1%;  Pred. No. 0.72;
  Matches  109;  Conservative    0;  Mismatches   89;  Indels    0;  Gaps    0;

Qy     209 ACCTGCGTTTGCTATAGGTGAAAATCCGATTTATTTGCCATAATGGAGACCCGTTTTTAT 268
           |||| | ||      |    ||  |||   | |||| ||||   | | | |  ||||| |
Db    1450 ACCTTCATTAATCTCAATCCAACTTCCTCATCATTTTCCATTTCGAATATCTCTTTTTCT 1509

Qy     269 TGAATATCGTCAGTTTCCAATTAAAGTCATTTCGAGCCATACTGAATAATTTTGTAATCT 328
           || | |    |  |||    ||    |   ||        |  |    | ||||  |   |
Db    1510 TGCACAGAAGCCTTTTTTCGTTTTTTTTTATTGATTTATTTTTACGGATTTTTAGATAAT 1569

Qy     329 AACAACAGATGCAATAGTTTAAATAAACTTATACTTAACTTTTAAAAAAAAAAAAAAAAA 388
             | |      || | | |||||||| ||||  | |   |  ||||||||||||||||||
Db    1570 GCACAGATGCCTCATTGCTCAAATAAATTTATTTTAATTGTCGAAAAAAAAAAAAAAAAA 1629

Qy     389 AAAAAAAAAAAAAAAAA 406
           |||||||||||||||||
Db    1630 AAAAAAAAAAAAAAAAA 1647
```

```
RESULT 11
AY121632
LOCUS       AY121632               1717 bp    mRNA    linear   INV 18-JUN-2002
DEFINITION  Drosophila melanogaster GM01935 full insert cDNA.
ACCESSION   AY121632
VERSION     AY121632.1  GI:21464311
KEYWORDS    FLI_CDNA.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1717)
  AUTHORS   Stapleton,M., Brokstein,P., Hong,L., Agbayani,A., Carlson,J.,
            Champe,M., Chavez,C., Dorsett,V., Dresnek,D., Farfan,D., Frise,E.,
            George,R., Gonzalez,M., Guarin,H., Kronmiller,B., Li,P., Liao,G.,
            Miranda,A., Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
            Patel,S., Phouanenavong,S., Wan,K., Yu,C., Lewis,S.E., Rubin,G.M.
            and Celniker,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-JUN-2002) Berkeley Drosophila Genome Project,
            Lawrence Berkeley National Laboratory, One Cyclotron Road,
            Berkeley, CA 94720, USA
COMMENT     Sequence submitted by:
            Berkeley Drosophila Genome Project
             Lawrence Berkeley National Laboratory
            Berkeley, CA 94720
             This clone was sequenced as part of a high-throughput process to
            sequence clones from Drosophila Gene Collection 1 (Rubin et al.,
            Science 2000). The sequence has been subjected to integrity checks
            for sequence accuracy, presence of a polyA tail and contiguity
            within 100 kb in the genome. Thus we believe the sequence to
            reflect accurately this particular cDNA clone. However, there are
            artifacts associated with the generation of cDNA clones that may
            have not been detected in our initial analyses such as internal
            priming, priming from contaminating genomic DNA, retained introns
            due to reverse transcription of unspliced precursor RNAs, and
            reverse transcriptase errors that result in single base changes.
            For further information about this sequence, including its location
            and relationship to other sequences, please visit our Web site
            (http://fruitfly.berkeley.edu) or send email to
            cdna@fruitfly.berkeley.edu.
FEATURES             Location/Qualifiers
     source          1. .1717
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
     gene            <1. .1717
                     /gene="RpI1"
                     /db_xref="FLYBASE:FBgn0019938"
     CDS             <1. .1629
                     /gene="RpI1"
                     /note="Longest ORF"
                     /codon_start=1
                     /product="GM01935p"
                     /protein_id="AAM51959.1"
                     /db_xref="GI:21464312"
                     /db_xref="FLYBASE:FBgn0019938"
                     /translation="LWKKADAEDKALYRKKYARCPDPTVAVYKQDLYYGSVSERTRKL
                     ITDYAKRKPALKETIADIMRVKTIKSLAAPGEPVGLIAAQSIGEPSTQMTLNTFHFAG
                     RGEMNVTLGIPRLREILMLASSNIKTPSMDIPIKPGQQHQAEKLRINLNSVTLANLLE
                     YVHVSTGLTLDPERSYEYDMRFQFLPREVYKEDYGVRPKHIIKYMHQTFFKQLIRAIL
                     KVSNASRTTKIVVIDDKKDADKDDDNDLDNGDEVGRSKAKANDDDSSDDNDDDDATGV
                     KLKQRKTDEKDYDDPDDVEELHDANDDDDEAEDEDDEEKGQDGNDNDGDDKAVERLLS
                     NDMVKAYTYDKENHLWCQVKLNLSVRYQKPDLTSIIRELAGKSVVHQVQHIKRAIIYK
                     GNDDDQLLKTDGINIGEMFQHNKILDLNRLYSNDIHAIARTYGIEAASQVIVKEVSNV
                     FKVYGITVDRRHLSLIADYMTFDGTFQPLSRKGMEHSSSPLQQMSFESSLQFLKSAAG
                     FGRADELSSPSSRLMVGLPVRNGTGAFELLTKIC"
BASE COUNT      532 a    369 c    432 g    384 t
ORIGIN
```

```
  Query Match            13.7%;  Score 55.6;  DB 3;  Length 1717;
  Best Local Similarity  82.1%;  Pred. No. 0.72;
  Matches   64;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy     329 AACAACAGATGCAATAGTTTAAATAAACTTATACTTAACTTTTAAAAAAAAAAAAAAAAA 388
           || |  |  || |||  |||||||||||| || | | ||| | |||||||||||||||||
Db    1638 AATATAAATTGAAATTTTTTAAATAAACATACATTAAACGTGTAAAAAAAAAAAAAAAAA 1697

Qy     389 AAAAAAAAAAAAAAAAA 406
           |||||||||||||||||
Db    1698 AAAAAAAAAAAAAAAAA 1715
```

```
RESULT 12
AF146568
LOCUS       AF146568               3024 bp    mRNA    linear   PRI 24-OCT-1999
DEFINITION  Homo sapiens MIL1 protein mRNA, complete cds; nuclear gene for
            mitochondrial product.
ACCESSION   AF146568
VERSION     AF146568.1  GI:6103438
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3024)
  AUTHORS   Zemskova,M.Y., Lilly,M. and Escher,A.P.
  TITLE     Mil1, a novel human gene encoding mitochondria located protein
            promoting cell survival
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 3024)
  AUTHORS   Zemskova,M.Y. and Escher,A.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-APR-1999) Center for Molecular Biology and Gene
            Therapy, Loma Linda University, 11085 Campus Street, Loma Linda, CA
            92350, USA
FEATURES             Location/Qualifiers
     source          1. .3024
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_line="glioma LNZ308"
     5'UTR           1. .71
     CDS             72. .1232
                     /function="involved in promoting cell survival"
                     /codon_start=1
                     /product="MIL1 protein"
                     /protein_id="AAF03602.1"
                     /db_xref="GI:6103439"
                     /translation="MEDCLAHLGEKVSQELKEPLHKALQMLLSQPVTYQAFRECTLET
                     TVHASGWNKILVPLVLLRQMLLELTRLGQEPLSALLQFGVTYLEDYSAEYIIQQGGWG
                     TVFSLESEEEEYPGITAEDSNDIYILPSDNSGQVSPPESPTVTTSWQSESLPVSLSAS
                     QSWHTESLPVSLGPESWQQIAMDPEEVKSLDSNGAGEKSENNSSNSDIVHVEKEEVPE
                     GMEEAAVASVVLPARELQEALPEAPAPLLPHITATSLLGTREPDTEVITVEKSSPATS
                     LFVELDEEEVKAATTEPTEVEEVVPALEPTETLLSEKEINAREESLVEELSPASEKKP
                     VPPSEGKSRLSPAGEMKPMPLSEGKSILLFGGAAAVAILAVAIGVALALRKK"
```

```
     3'UTR           1233. .3024
BASE COUNT      850 a    701 c    728 g    745 t
ORIGIN

  Query Match            13.6%;  Score 55.4;  DB 9;  Length 3024;
  Best Local Similarity  61.4%;  Pred. No. 0.73;
  Matches   89;  Conservative    0;  Mismatches   56;  Indels    0;  Gaps    0;

Qy     262 TTTTTATTGAATATCGTCAGTTTCCAATTAAAGTCATTTCGAGCCATACTGAATAATTTT 321
           ||| | ||  | ||      |||| ||| |||      |  ||| ||| | || ||||
Db    2859 TTTCTTTTTGAAATTTATGTTTTCAAATAAAATCTCCCTAAAGCAATATTTAAAAATTGG 2918

Qy     322 GTAATCTAACAACAGATGCAATAGTTTAAATAAACTTATACTTAACTTTTAAAAAAAAAA 381
             ||   || || | |   || |    ||| |||   | |   ||     ||||||||||
Db    2919 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2978

Qy     382 AAAAAAAAAAAAAAAAAAAAAAA 406
           |||||||||||||||||||||||
Db    2979 AAAAAAAAAAAAAAAAAAAAAAA 3003


RESULT 13
AC116978
LOCUS       AC116978               27785 bp    DNA     linear   HTG 31-MAY-2002
DEFINITION  Dictyostelium discoideum chromosome 2 map 4846697-4874480 strain
            AX4, *** SEQUENCING IN PROGRESS ***, in ordered pieces.
ACCESSION   AC116978
VERSION     AC116978.1  GI:20042933
KEYWORDS    HTG; HTGS_PHASE2.
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 27785)
  AUTHORS   Gloeckner,G., Eichinger,L., Szafranski,K., Pachebat,J., Dear,P.,
            Lehmann,R., Baumgart,C., Parra,G., April,J.F., Guigo,R., Kumpf,K.,
            Tunggal,B., Cox,E., Quail,M.A., Platzer,M., Rosenthal,A. and
            Noegel,A.A.
  TITLE     Sequence and Analysis of Chromosome 2 of Dictyostelium
  JOURNAL   Unpublished
  REMARK    The Dictyostelium Genome Sequencing Consortium
REFERENCE   2  (bases 1 to 27785)
  AUTHORS   Baumgart,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-APR-2002) Genome Analysis, Institute of Molecular
            Biotechnology, Beutenberstr. 11, Jena 07745, Germany
REFERENCE   3  (bases 1 to 27785)
  AUTHORS   Baumgart,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-MAY-2002) Genome Analysis, Institute of Molecular
            Biotechnology, Beutenberstr. 11, Jena 07745, Germany
COMMENT     CDS predictions from GeneID may contain errors. Further Information
            is available from IMB Jena, Department of Genome Analysis
            (http://genome.imb-jena.de/dictyostelium/)
            and the Univerity Colonge, Institute for Biochemistry I
            (http://www.uni-koeln.de/dictyostelium/project.shtml
            Funding
            Agency : Deutsche Forschungsgemeinschaft (DFG).
            * NOTE: This is a 'working draft' sequence.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
FEATURES             Location/Qualifiers
     source          1. .27785
                     /organism="Dictyostelium discoideum"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /chromosome="2"
                     /map="4846697-4874480"
     CDS             678. .1491
                     /note="ORF_ID:dd_01248"
                     /pseudo
                     /codon_start=1
     CDS             join(2445. .2467,2559. .3372)
                     /note="ORF_ID:dd_01250"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="AAM45259.1"
                     /db_xref="GI:21281372"
                     /translation="MPLALTSRKSLKDNEESLKTNLASIEKSLGEPFSIEFDFEDIIT
                     KCNDSWVTNSCGSIFYKDVVGNLAKNMVKVGADPLIKEAFLAAVPNKKFVFVCADEKL
                     ESYWKYQFENGDCRVLFRPKICNTNDVNTFKLETIIPTQGVYTLMTRLNIKTNEQKMK
                     DNLSAVKKALKSSSDWSIDESSLETVYPKVADDLKNSFGHIFAGIVEKVAANLAKRCA
                     DEMILEAVQEATTNHTIVIKHDANLNAHWSWSFESGNIVITFKSVTNTNDVQTFDFIK
                     LL"
     CDS             complement(3680. .4780)
                     /note="ORF_ID:dd_01251"
                     /pseudo
                     /codon_start=1
     CDS             join(6336. .6630,6711. .11386,11471. .14377)
                     /note="ORF_ID:dd_00705"
                     /codon_start=1
                     /product="Hypothetical 178.4 kDa protein"
                     /protein_id="AAM45260.1"
                     /db_xref="GI:21281373"
                     /translation="MWTARAAKFLFKKICGEYFMYELEKDQFKIAFKKGTINLNNLEL
                     NIKKINKDLYGVSPLMLASGYIGNLTTMIPYTSLLDTPSEFNISNLELTFVSSNDYEK
                     YVLEGLKQLKYLDHSSTNKTNNDDNNNNKNENVNYYSDEDEEIDNIETKKQFGIQGLE
                     ALTEIVKKLIDKVSARIDNCTITIVEYDQSRNKSIIIVFLIESIEYRASSTGDPPNYF
                     NNNNNNNNNNNNTSSQEKIDYFLKSIVLKNHSILITDKYTGNILLEEKFYNNPIFNI
                     IPKSKFDIIFSNKVDSSNVNNNNNNNSSNNVNNVNNNDIISLKIRRDDKPSPTTQPI
                     KITNTNTNTNNNELPMFLVDCFFKQWTIVLSPYHLSIIDSFISNFNKPIQKSTTTTTT
                     TTTTTNSIPNIKTKRQSISKNSSNQLFNLNLKWQSIKIYIHKDRKENYYNEFIPMMVI
                     DKLNPPQPPKRDYLALKLFDINYSLKQQNQIKKMNLTISKLLITNYNVNSTCGELINL
                     EIIKMPISSNSPNSSSSSIVSNLSSNSLNKLVNSNTNNNSNQSYSPPLQIQQQQQQQQ
                     QQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQQINEFNCNIIFEESNMILDIQIKDELNLIMD
                     PFLLLFFNSFTSEMKKVEKEIKEEKRLHKLYLNRIEKQNIKNNILKKQKDQQLRLQR
                     KELQKQKQLQHQRQHQSSDYYYPHEVDYYDSDEDGDEYYYHEGDSFKNNQINNNNNNI
                     NNENNKENLTQSNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN
                     NNNNNNDNNDEEQDEIESSSIQFSLQSSVIRFKVKFPSFEFIPNSLVIDLKDQKKFK
                     STFFEIQKNLKLLTNQLEISFTQCKINSNIIKDIPDVKWFIDFENLSLLLKDINTNTS
                     GNNYNNNNNNNNNNNNNNNNNNNNNNNNSNNNNNINSIINIITKNDSDSFNPVPIEITVRG
                     LLNNKNTGNDSNKDVDDHQDNGIEKDVEQGQEDKEDIDNIDYSGGGSGANIITNDDSY
                     DDIDDDSDIFEMTNDDIDMYDIEGEDNVKIPQNIPQPQLPSPPTPTPIPPQQQQQQPQ
                     PPSQQQQQQQKQHQQGNTSPNFSHFNPFSKSSEKHKGPFSTFVSTSEGVFDSWPDEAE
                     QEEINIFRHSSIESSRYAIYLIIPFLQLKLTKSQYQSLIQLSQSYSKLIDYTTIPKSS
                     VSSFINNNNNNNNNNNNNNNNNNNNSNNNNNINSIINIITKNDSDSFNPVPIEITVRG
                     LLNNKNTGNDSNKDVDDHQDNGIEKDVEQGQEDKEDIDNIDYSGGGSGANIITNDDSY
                     ILPRSKNGSPQVKSFNNNNNNNNNDYNNYSTATPSSVVSPIEEVSLFSLMVVINSGS
                     FIIQEDAPKSSFSNSRNNSSNNIKGQQYQPPPPPPPQQQPWYQYELSFEKFEIFNVSQ
                     YYGRDISFVSSMVSKVKLTELNLTSKDNKPLPIIFKTISQERNSTDQMISVTISFNRD
                     TTGFLDDNIHGNIYSFVFKGVTIYHKLNTQWYSRISNFFNDSTNNNNNNNNNNNNNN
                     NNNNNNTTSNENPKDKEEEKEQEEEIEEEKRRKEEEIKKELLKNREEEEVNNTQSRYF
                     INFVDTSISYTPSDNLSSAAVLFFSDVKLKIPTSGFFSCIGQNSYIYVMDDIQFLNLN
                     KPNSKSSSSNSILNYWKYLGYVPVVHLDYIEFDITNSSDNDHYPSLKIDYNHNLLSIN
                     ACSDSFYILNQLINNFLGFDLPIAEVADLIITQGIGSELNLKESIDILINEDTFRQTA
                     EQLQKQQQQQQQQSYSSNNINLSSSPSGGPTINSSNNNNNNNSNLSSSPLVPGAINFE
                     FEEVLRGDGNNDFSDDDGSSSVYEYYSPLDMMSSNSKLSETESNCSENQMINEIHDTI
                     SSHTGSFSNLPIMGSIMIEDYDHDKLETYPPSSTTQMYPHRDFYNDIDPPSIQLQQQP
                     TQSSSYQPQQHQQQQQQQQQQQQQFNSVNDGSVTWIDSTNQNKSIEPVENYISNPDED
                     PSEDCPLPSHYPKSASKMSFRKMNLSIKIFKGLDWERESTIIEPSSPSSLAGGKEKQQ
                     SQQQQQPQHKSGRDQNCFVEFILSNVNIRIDRFDDNELYAKRMSLHISDVTIMDHIPT
                     SLWNKFLCSDTNVTRYQQSPMIKILFETVRPDLARPLLQENRLKVLILPIRFNIDQDT
                     VSFMVNYFSYKPPSSPSPIQPSIVIESPRSTNNNNNNNTSVSPQNVSPKINNITTTTT
                     TTSITTNTTTTVTSSISPNSSIISPLSTSQTFTPGVSSSSLLNSSGGNSISSVTSSSS
                     SSINTTSSNVTPIPIKKNSSSSSSSSQTISRSIPEDKAQTTDKESQEITYFQSVEILP
                     IRLKVDYKPKKVDYHSLTTGNYAELLNLLPLEGAIFNLSKLKLNGVGGWSALFNQIAK
                     IWVPHIIQTQLLGYVSGVRGLNSLVHIGEGLANLILPYEQYKKDGNILKGIKRGTTT
                     FLKNLTVETLSMGAKVAVGTQGLLETADNVLSSPKQQHSQHHSQHQHSQYHHQSSSSS
                     SSTQKHHSSSSHGRFKNHQSILHSFNSPTQQLSPASIQSLNHHLQNHHGISKFSDQPN
                     DTKEGIQHAYESVSRELKSAAHTIIAIPMKEYHQKGTKGYVKSIVQAVPIAIIRPMIG
                     ITEGVSKTLLGVRNQIDPNKKLEMDNKYK"
     CDS             complement(join(14916. .15729,15822. .15844))
                     /note="ORF_ID:dd_00706"
                     /codon_start=1
                     /product="hypothetical protein"
                     /protein_id="AAM45261.1"
                     /db_xref="GI:21281374"
```

```
                /translation="MPLALTSRKSLKDNEESLKTNLASIEKSLGEPFSIEFDFEDIIT
                KCNDSWVTNSCGSIFYKDVVGNLAKNMVKVGADPLIKEAFLAAVPNKKFVFVCADEKL
                ESYWKYQFENGDCRILFRPKICNTNDVNTFKLETIIPTQGVYTLMTRLNIKTNEQKMK
                DNLSAVKKALKSSSDWSIDESSLETVYPKVADDLKNSFGHIFAGIVEKVAANLAKRCA
                DEMILEAVQEATTNHTIVIKHDANLNAHWSWSFESGNIVITFKSVTNTNDVQTFDFIK
                LL"
CDS             complement(join(16638. .17451,17546. .17568))
                /note="ORF_ID:dd_00707"
                /codon_start=1
                /product="hypothetical protein"
                /protein_id="AAM45262.1"
                /db_xref="GI:21281375"
                /translation="MPLALTSRKSLKDNEESLKTNLASIEKSLGEPFSIEFDFEDIIT
                KCNDSWVTNSCGSIFYKDVVGNLAKNMVKVGADSLIKEAFLAAVPNKKFVFVCADEKL
                ESYWKYQFENGDCRILFRPKICNTNDVNTFKLETIIPSQGVYTLMTRLNIKTNEQKMK
                DNLSAVKKALKSSSDWSIDESSLETVYPKVADDLKNSFGHIFAGIVEKVAANLAKRCA
                DEMILEAVQEATTNHTIVIKHDANLNAHWSWSFESGNIVITFKSVTNTNDVQTFDFIK
                LL"
CDS             complement(join(18447. .19260,19351. .19373))
                /note="ORF_ID:dd_00708"
                /codon_start=1
                /product="hypothetical protein"
                /protein_id="AAM45263.1"
                /db_xref="GI:21281376"
                /translation="MPLALTSRKSLKDNEESLKTNLASIEKSLGEPFSIEFDFEDIIT
                KCNDSWVTNSCGSIFYKDVVGNLAKNMVKVGADSLIKEAFLAAVPNKKFVFVCADEKL
                ESYWKYQFENGDCRILFRPKICNTNDVNTFKLETIIPSQGVYTLMTRLNIKTNEQKMK
                DNLSAVKKALKSSSDWSIDELSLETVYPKVADDLKNSFGHIFAGIVEKVAANLAKRCA
                DEMILEAVQEATTNHTIVIKHDANLNAHWSWSFESGN1VITFKSVTNTNDVQTFDFIK
                LL"
CDS             19866. .21218
                /note="ORF_ID:dd_00709"
                /codon_start=1
                /product="hypothetical protein"
                /protein_id="AAM45264.1"
                /db_xref="GI:21281377"
                /translation="MILEFKLLYEKNEIVTFLNLTSLIEYENKEYIRTLIYKGNESIN
                EKCFLPNNKNGVLNKIVFKEQSISNIASCLIPFGVEIIEFGISLTFNNNILVSTTPLF
                LLPSTIKSVKNVVINETDYSKGNDYLIPNNIKEITFLKCPKISKDTIKFPSTLESITF
                INEWNNGDSILKIGDIPNGIKNLVLNSYREGFDISVLPNSIINLEVILKCNKLVKNVI
                PINVKSLTMMYYINENENENNNNTNKLEITHESLPPSLINLECYSKDMKFLPNSIPSTV
                ENLQIYDLQIPLTPSILSPSNSIKKFEIAFDLIEFTCKRLNLNLNYYIGSLLQCLTSL
                IEFSIIGYFNSKIDTNILPNGLKIFNLLENINFNQPFDQDNILPSTLTHLILNNNYSI
                ENLLKIKLPISLTFLSLSKNLKNNFNINNFNNFNNNNNNNNNNNNNNNKNLKIHFRD
                "
CDS             complement(join(21354. .23054,23166. .23375))
                /note="ORF_ID:dd_00710"
                /codon_start=1
                /product="hypothetical protein"
                /protein_id="AAM45265.1"
                /db_xref="GI:21281378"
                /translation="MKDIWKSARKGDLAKIKKYIPGTKFKVDELDKENRSILHHAIDG
                DQLQMVNYLLEQGANINLKDPKSGDTPILMSLKKINAIDKQEDIAILLVNKGADLDVE
                SRERKNVERYVDACPKRFQIKYKDAITQKFKNKYPPKAIAPASPATQVLLAQTAGLPP
                PSPKPNPPPPQQQENTATSSIISTVEPRPYVNSNGETVVRSTNEMDLILQESLLGDDF
                DKEQDLNIIWMEVAFRASIVSIRSEKVRYDDITKATDQCGNILIHLLSIIGNSETTSF
                METVKNQEAVNSIQERVKNVKEGIKLRLFSRIDHSELIRLVAALIGSLHWIYNCWSEV
                PQDDINRSIAECALSCRVAMNSVTTHTPFQLGVLTNIAKLLKNISTKVFITRTPEDAK
                KLSDSCFVISSTFKSLMVLSVVNDVKQKSAENIISLGRSIAEQLSILKLQSTSILPSR
                VGQNLSSKEERELNESASNQLRSSIDYYISNNKSDTTTTTTTTTTTNSENNNITDLI
                SSPKRLPKEDKTLPLLANIEKSISMLMGFTTEDRNIQMAIPVHMIAKDIHELRTILLE
                YIKEQSDSLSEEFKDQITNAINCSAHFSTQQLFSVSALICHNRITDISQLGYSSRNIA
                ACCCILIDSFCF"
CDS             join(24733. .24739,24851. .27393)
                /note="ORF_ID:dd_00704"
                /codon_start=1
                /product="hypothetical protein"
                /protein_id="AAM45266.1"
                /db_xref="GI:21281379"
                /translation="MSVLSVGQSKNGQLGHDSIKECLEPTVIEGLLGKRIIQTVSGEA
                HSLAVSRFGDVYSWGRSKEGQLGIGQGYGGSDKVMFVAKPTLIKSLQHERIIKVACGN
                FHSLALTDMGKVYEWGQLHRLDESQSSLDIQSTNGLIEMPRLSSQRIIEASVSQYLSG
                EKKAYDANGGTKPKDDDEDKEASNSAQSKSKHLGKIIDLNQTIPMQIQDIPEEVVDVS
```

```
                GGWAFSAAVTKSGRLYTWGFNEKGQLGLGNRWFHSTPQLVKTLIDVNIVSVVCGRQHI
                CAITDQGEVYSWGLGVFGQLGHGNVKSYLHPKKIQQFVELNERIAQVACGSNFTMVRS
                VQGLLYAFGHGEYGQLGSTEETQHLDFGGRDNHFKYSIPIVVKSLETKKIKNVACGHL
                HTIVVTDENEVYQWGWGSSGALGLGNKRFQLVPQLITSLSGEEIASITAGEKHTIVVR

  Query Match             13.6%;  Score 55.4;  DB 2;  Length 27785;
  Best Local Similarity   58.0%;  Pred. No. 0.56;
  Matches   98;  Conservative    0;  Mismatches   71;  Indels    0;  Gaps    0;

Qy     238 TTTATTTGCCATAATGGAGACCCGTTTTTATTGAATATCGTCAGTTTCCAATTAAAGTCA 297
           | | |||   ||  |    |    ||| | ||   |||  | | |||  | ||||| | |
Db    6114 TATTTTTTTTATTTTTTTAATTTTTTTATTTTTTTTATTTTTATTTTAAATTTAAATTTA 6173

Qy     298 TTTCGAGCCATACTGAATAATTTTGTAATCTAACAACAGATGCAATAGTTTAAATAAACT 357
           | |      | |  ||   ||   | |   | ||||| |   |||| || ||| |||
Db    6174 CTACATTTTTTCCACAACTTTTAAATTAATAATCAACATAACAAATAATTAAAACAAAAC 6233

Qy     358 TATACTTAACTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 406
           | ||  || |   ||||||||||||| ||||||| |||||| ||||||
Db    6234 AAAACAAAATTAAAAAAAAAAAAAAAATAAAAAAATAAAAAACAAAAAA 6282


RESULT 14
AY069596
LOCUS       AY069596               1432 bp    mRNA    linear   INV 17-DEC-2001
DEFINITION  Drosophila melanogaster LD34980 full length cDNA.
ACCESSION   AY069596
VERSION     AY069596.1  GI:17862527
KEYWORDS    FLI_CDNA.
SOURCE      Drosophila melanogaster.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1432)
  AUTHORS   Stapleton,M., Brokstein,P., Hong,L., Agbayani,A., Carlson,J.,
            Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George,R.,
            Gonzalez,M., Guarin,H., Li,P., Liao,G., Miranda,A., Mungall,C.J.,
            Nunoo,J., Pacleb,J., Paragas,V., Park,S., Phouanenavong,S., Wan,K.,
            Yu,C., Lewis,S.E., Rubin,G.M. and Celniker,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-DEC-2001) Berkeley Drosophila Genome Project,
            Lawrence Berkeley National Laboratory, One Cyclotron Road,
            Berkeley, CA 94720, USA
COMMENT     Sequence submitted by:
            Berkeley Drosophila Genome Project
            Lawrence Berkeley National Laboratory
            Berkeley, CA 94720
            This clone was sequenced as part of a high-throughput process to
            sequence clones from Drosophila Gene Collection 1 (Rubin et al.,
            Science 2000). The sequence has been subjected to integrity checks
            for sequence accuracy, presence of a polyA tail and contiguity
            within 100 kb in the genome. Thus we believe the sequence to
            reflect accurately this particular cDNA clone.  However, there are
            artifacts associated with the generation of cDNA clones that may
            have not been detected in our initial analyses such as internal
            priming, priming from contaminating genomic DNA, retained introns
            due to reverse transcription of unspliced precursor RNAs, and
            reverse transcriptase errors that result in single base changes.
            For further information about this sequence, including its location
            and relationship to other sequences, please visit our Web site
            (http://fruitfly.berkeley.edu) or send email to
            cdna@fruitfly.berkeley.edu.
FEATURES             Location/Qualifiers
     source          1. .1432
                     /organism="Drosophila melanogaster"
                     /strain="y; cn bw sp"
                     /db_xref="taxon:7227"
                     /map="7A4-7A4"
     gene            1. .1432
                     /gene="CG1643"
                     /note="alignment with genomic scaffold AE003440"
                     /db_xref="FLYBASE:FBgn0029943"
```

```
     CDS             69. .878
                     /gene="CG1643"
                     /note="Longest ORF"
                     /codon_start=1
                     /product="LD34980p"
                     /protein_id="AAL39741.1"
                     /db_xref="GI:17862528"
                     /db_xref="FLYBASE:FBgn0029943"
                     /translation="MAHDREVLRMIWEGQIGICFQADRDEIVGIKPEPFYLMISRLSY
                     LPLVTDKVRKYFSRYISAEHQDGAVWFDFNGTPLRLHYPIGVLYDLLHPEEDSTPWCL
                     TIHFSKFPEDMLVKLNSKELLESHYMSCLKEADVLKHRGLVISAMQKKDHNQLWLGLV
                     NEKFDQFWAVNRRLMEPYGDLESFKNIPLRIYTDDDFTYTQKLISPISVGGQKKSLAD
                     LMAELSTPVRRAVGCRTHGIDLHEETQLQWMSEHLSYPDNFLHLSVDYKDV"
BASE COUNT      441 a    312 c    301 g    378 t
ORIGIN

  Query Match             13.5%;  Score 55;  DB 3;  Length 1432;
  Best Local Similarity   66.4%;  Pred. No. 0.95;
  Matches   79;  Conservative    0;  Mismatches   40;  Indels    0;  Gaps    0;

Qy     288 ATTAAAGTCATTTCGAGCCATACTGAATAATTTTGTAATCTAACAACAGATGCAATAGTT 347
           | |||| | |  ||  |||| ||| |  | |  || || |||         |    | |
Db    1302 ACTAAAATGACATCACGCCACACTAATGATTAATGGAAACTATTTTATATCGACCCAATC 1361

Qy     348 TAAATAAACTTATACTTAACTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 406
           | ||| |  |||||  || || ||||||||||||||||||||||||||||||||||||
Db    1362 AATATATATATATACGAAAGTTATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1420


RESULT 15
DDIDPP5A
LOCUS       DDIDPP5A                1805 bp    DNA     linear   INV 11-MAR-1994
DEFINITION  Dictyostelium purpureum (Dpp5) DNA sequence, repeat region.
ACCESSION   L05617
VERSION     L05617.1  GI:167774
KEYWORDS    repeat region.
SOURCE      Dictyostelium purpureum (strain DPA) DNA.
  ORGANISM  Dictyostelium purpureum
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 1805)
  AUTHORS   Kiyosawa,H., Hughes,J.E., Podgorski,G.J. and Welker,D.L.
  TITLE     Small circular plasmids of the eukaryote Dictyostelium purpureum
            define two novel plasmid families
  JOURNAL   Plasmid 30 (2), 106-118 (1993)
  MEDLINE   94052367
   PUBMED   8234483
FEATURES             Location/Qualifiers
     source          1. .1805
                     /organism="Dictyostelium purpureum"
                     /strain="DPA"
                     /db_xref="taxon:5786"
     repeat_region   320. .477
BASE COUNT      783 a    159 c    202 g    661 t
ORIGIN

  Query Match             13.5%;  Score 55;  DB 3;  Length 1805;
  Best Local Similarity   61.5%;  Pred. No. 0.92;
  Matches   88;  Conservative    0;  Mismatches   55;  Indels    0;  Gaps    0;

Qy     264 TTTATTGAATATCGTCAGTTTCCAATTAAAGTCATTTCGAGCCATACTGAATAATTTTGT 323
           || ||| ||| | |  |  |   ||| |||   | ||  |   |||    ||||  ||
Db    1154 TTAATTCAATTTTGATATATAAAAATAAAAAATAGTTATATTGATAAACTATAAAATTTA 1213

Qy     324 AATCTAACAACAGATGCAATAGTTTAAATAAACTTATACTTAACTTTTAAAAAAAAAAAA 383
           ||   || || | |   |||  || ||| |||  ||||       | || |||| |||||
Db    1214 AAAAAAAAAAAAAAAAAAATTTTTAAAAAAAAAATATATGCCGGATATATAAAATAAAAA 1273

Qy     384 AAAAAAAAAAAAAAAAAAAAAA 406
           ||||||||||| ||||||||||
Db    1274 AAAAAAAAAATAAAAAAAAAA 1296


Search completed: May  5, 2003, 19:53:08

Job time : 2057 secs
```

GenCore version 5.1.5

OM nucleic - nucleic search, using sw model

Run on: May 5, 2003, 18:00:58 ; Search time 230 Seconds
(without alignments)
3975.267 Million cell updates/sec

Title: US-09-829-481-3
Perfect score: 406
Sequence: 1 ctctactacaatcactaagt.....aaaaaaaaaaaaaaaaaaaa 406

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	406	100.0	406	24	AAH77209	Arthropod defensin
	2	76	18.7	386	24	AAH77210	Arthropod defensin
	3	59.2	14.6	230	23	ABV57465	Human prostate exp
	4	59.2	14.6	236	23	ABV56953	Human prostate exp
	5	58	14.3	574	23	ABV57502	Human prostate exp
	6	57	14.0	497	23	ABV58273	Human prostate exp
	7	56	13.8	381	23	ABV58662	Human prostate exp
c	8	55.8	13.7	17294	24	ABL32986	Human immune syste
	9	55.2	13.6	325	22	AAS29049	cDNA encoding for
	10	55.2	13.6	621	22	AAS29124	cDNA encoding for
	11	54.6	13.4	351	24	AAH77211	Arthropod defensin
c	12	54.6	13.4	425	22	AAS60450	Human cancer agent
	13	54.6	13.4	2736	24	ABQ54618	Human ovarian anti
c	14	54.4	13.4	372	23	ABV37528	Human prostate exp
c	15	54.4	13.4	447	23	ABV10351	Human prostate exp
c	16	54.2	13.3	266	23	ABV07596	Human prostate exp
	17	54.2	13.3	1327	24	AAD24775	Glycine max ankyri
	18	53.8	13.3	309	23	ABV44994	Human prostate exp
c	19	53.8	13.3	431	23	ABV04399	Human prostate exp
	20	53.8	13.3	461	24	AAH77208	Arthropod defensin
c	21	53.8	13.3	6290	24	ABL33047	Human immune syste
	22	53.6	13.2	154	22	AAL23561	Human breast cance
	23	53.6	13.2	193	22	AAL14699	Human breast cance
	24	53.6	13.2	1812	21	AAA16649	Human secreted pro
c	25	53.4	13.2	6189	22	AAS46600	Tumour suppressor
c	26	53.4	13.2	6189	24	ABL34593	Human metastasis a
c	27	53.2	13.1	300	22	AAH70049	Human cervical can
	28	53	13.1	959	24	ABK63920	cDNA encoding huma
	29	52.8	13.0	3581	15	AAQ44391	Sequence of murine
	30	52.6	13.0	3293	22	AAS25944	Human cDNA encodin
c	31	52.6	13.0	5198	24	ABL32279	Human immune syste
c	32	52.4	12.9	165	22	AAI84674	Human polynucleoti
	33	52.4	12.9	958	22	AAC92233	Apoaequorin-encodi
	34	52.4	12.9	958	24	AAD22188	Aequorea victoria
	35	52.4	12.9	959	9	AAN81534	pAQ440 aequorin ge
	36	52.4	12.9	1994	22	AAH48294	Honeybee alpha-glu
	37	52.4	12.9	2175	21	AAC79978	Human secreted pro
c	38	52.2	12.9	368	22	AAS60046	Human cancer agent
	39	52.2	12.9	426	22	AAI88844	Human polynucleoti
	40	52.2	12.9	441	22	AAI91749	Human polynucleoti
	41	52.2	12.9	504	22	AAI85190	Human polynucleoti
	42	52.2	12.9	612	22	AAH71471	Human cervical can
	43	52.2	12.9	815	21	AAC79968	Human secreted pro
	44	52.2	12.9	1387	22	AAS29132	cDNA encoding for
	45	52.2	12.9	3957	22	ABA06411	Human cDNA SEQ ID

ALIGNMENTS

RESULT 1
AAH77209
ID AAH77209 standard; cDNA; 406 BP.
XX
AC AAH77209;
XX
DT 30-JAN-2002 (first entry)
XX
DE Arthropod defensin cDNA from Vaejovis carolinianus clone iks1c.pk0.
XX
KW Arthropod; defensin; iks1c.pk0; antimicrobial; microbial pathogen;
KW gene therapy; ss.
XX
OS Vaejovis carolinianus.
XX
FH Key Location/Qualifiers
FT CDS 44..229
FT /*tag= a
FT /product= "Defensin"
XX
PN EP1146052-A2.
XX
PD 17-OCT-2001.
XX
PF 12-APR-2001; 2001EP-0303488.
XX
PR 14-APR-2000; 2000US-197279P.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Presnail JK, Weng Z, Wong JF;
XX

```
DR   WPI; 2002-001056/01.
DR   P-PSDB; AAG77932.
XX
PT   New polynucleotides encoding Arthropod defensin polypeptides for
PT   improving the microbial resistance of plants and animals by gene
PT   therapy -
XX
PS   Claim 6; Page 18; 22pp; English.
XX
CC   The sequence represents a novel Arthropod defensin from V. carolinianus
CC   clone iks1c.pk0. The invention relates to novel polynucleotides encoding
CC   Arthropod defensin polypeptides. The polypeptides and polynucleotides of
CC   the invention have antimicrobial activity. The polynucleotides may be
CC   used to genetically engineer cells and organisms, especially plants, to
CC   alter their resistance to a wide range of microbial pathogens. The
CC   polynucleotides may also have a use in gene therapy.
XX
SQ   Sequence 406 BP; 143 A; 66 C; 67 G; 130 T; 0 other;

  Query Match             100.0%;  Score 406;  DB 24;  Length 406;
  Best Local Similarity   100.0%;  Pred. No. 3.1e-64;
  Matches  406;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CTCTACTACAATCACTAAGTTCTTTCTCCACTCAGCTTCAAGAATGAAATCCATAGCTAT 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CTCTACTACAATCACTAAGTTCTTTCTCCACTCAGCTTCAAGAATGAAATCCATAGCTAT 60

Qy      61 TATTTTCATCGTTCTTGTTGCCTTCTGTATTTTGGAGGATGGGATTGTAGAAGCTGGTTT 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TATTTTCATCGTTCTTGTTGCCTTCTGTATTTTGGAGGATGGGATTGTAGAAGCTGGTTT 120

Qy     121 TGGATGTCCCTTTAATGCAGGAAAATGCCATAGACATTGCAAAAGTATTCGTCGTAGAGG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 TGGATGTCCCTTTAATGCAGGAAAATGCCATAGACATTGCAAAAGTATTCGTCGTAGAGG 180

Qy     181 AGGCTTTTGCAGAGGAACTTTCAGGACAACCTGCGTTTGCTATAGGTGAAAATCCGATTT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 AGGCTTTTGCAGAGGAACTTTCAGGACAACCTGCGTTTGCTATAGGTGAAAATCCGATTT 240

Qy     241 ATTTGCCATAATGGAGACCCGTTTTTATTGAATATCGTCAGTTTCCAATTAAAGTCATTT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 ATTTGCCATAATGGAGACCCGTTTTTATTGAATATCGTCAGTTTCCAATTAAAGTCATTT 300

Qy     301 CGAGCCATACTGAATAATTTTGTAATCTAACAACAGATGCAATAGTTTAAATAAACTTAT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CGAGCCATACTGAATAATTTTGTAATCTAACAACAGATGCAATAGTTTAAATAAACTTAT 360

Qy     361 ACTTAACTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 406
          ||||||||||||||||||||||||||||||||||||||||||||||
Db     361 ACTTAACTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 406


RESULT 2
AAH77210
ID   AAH77210 standard; cDNA; 386 BP.
XX
AC   AAH77210;
XX
DT   30-JAN-2002  (first entry)
XX
DE   Arthropod defensin cDNA from Argiope sp. clone aot1c.pk0 #1.
XX
KW   Arthropod; defensin; aot1c.pk0; antimicrobial; microbial pathogen;
KW   gene therapy; ss.
XX
OS   Argiope sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             14..196
FT                   /*tag=  a
FT                   /product= "Defensin"
XX
PN   EP1146052-A2.
XX
PD   17-OCT-2001.
XX
PF   12-APR-2001; 2001EP-0303488.
XX
PR   14-APR-2000; 2000US-197279P.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Presnail JK,  Weng Z,  Wong JF;
XX
DR   WPI; 2002-001056/01.
DR   P-PSDB; AAG77933.
XX
PT   New polynucleotides encoding Arthropod defensin polypeptides for
PT   improving the microbial resistance of plants and animals by gene
PT   therapy -
XX
PS   Claim 6; Page 18; 22pp; English.
XX
CC   The sequence represents a novel Arthropod defensin from Ariope sp.
CC   clone aot1c.pk0. The invention relates to novel polynucleotides encoding
CC   Arthropod defensin polypeptides. The polypeptides and polynucleotides of
CC   the invention have antimicrobial activity. The polynucleotides may be
CC   used to genetically engineer cells and organisms, especially plants, to
CC   alter their resistance to a wide range of microbial pathogens. The
CC   polynucleotides may also have a use in gene therapy.
XX
SQ   Sequence 386 BP; 140 A; 81 C; 67 G; 94 T; 4 other;

  Query Match             18.7%;  Score 76;  DB 24;  Length 386;
  Best Local Similarity   54.4%;  Pred. No. 3.8e-05;
  Matches  178;  Conservative    0;  Mismatches  141;  Indels    8;  Gaps    1;

Qy      78 TTGCCTTCTGTATTTTGGAGGATGGGATTGTAGAAGCTGGTTTTGGATGTCCCTTTAATG 137
           | | | |||||  ||| |     | |   || ||||||||||| || || |||||  |
Db      45 TAGTCGTCTGTGCTTTTGCCACAGTGGCCGTGGAAGCTGGTTTCGGCTGCCCCTTCGACC 104

Qy     138 CAGGAAAATGCCATAGACATTGCAAAAGTATTCGTCGTAGAGGAGGCTTTTGCAGAGGAA 197
               | || || |  |||||||  || ||        || ||||| |  ||||
Db     105 AGATGCAGTGTCACAATCATTGCAGGAGCATCAAATACAGGGGAGGATACTGCACCAACT 164

Qy     198 CTTTCAGGACAACCTGCGTTTGCTATAGGTGAAAATCCGATTTATTTGCCATAATGGAGA 257
             |||| |   ||||||   || ||  | |||  | ||   |    |  || | ||||
Db     165 TATTCAAGCGCACCTGCAAGTGTTACGGATGATGACCCCCCTCCCCTCTCAGACAGGAGC 224

Qy     258 CCCGTTTTTATTGAATATCGTCAGTTTCCAATTAAAGTCATTTCGAGCCATACTGAATAA 317
           ||     ||         || |     ||| |||    | |    |  || | || |  |
Db     225 CCACACCTT--------TCATTGACATCCGATTCCGATTTTCCAAATGCAAATTGTAACA 276

Qy     318 TTTTGTAATCTAACAACAGATGCAATAGTTTAAATAAACTTATACTTAACTTTTAAAAAA 377
             |  | |  |     ||  |||  ||  || ||  || ||  | || || ||| |||||
Db     277 CATGATGAATTTGATGCAAGTGCCTTAAATTTAAATAAATTTGATTTTACATTTTAAAAA 336

Qy     378 AAAAAAAAAAAAAAAAAAAAAAAAAAAA 404
           |||||||||||||| ||||||||||||
Db     337 AAAAAAAAAAAAAAANAAAAAAAAAAAA 363


RESULT 3
ABV57465
ID   ABV57465 standard; cDNA; 230 BP.
XX
AC   ABV57465;
XX
DT   17-SEP-2002  (first entry)
XX
DE   Human prostate expression marker cDNA 57456.
XX
KW   Human; prostate cancer; cytostatic; carcinogen; pharmacodyanamic marker;
KW   pharmacogenomic marker; gene; ss.
XX
OS   Homo sapiens.
```

XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US05171.
XX
PR 17-FEB-2000; 2000US-183319P.
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
DR WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer -
XX
PS Claim 1; Page 11053; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (I) is also useful as a pharmacodyanamic or pharmacogenomic marker.
XX
SQ Sequence 230 BP; 101 A; 32 C; 45 G; 52 T; 0 other;

Query Match 14.6%; Score 59.2; DB 23; Length 230;
Best Local Similarity 79.5%; Pred. No. 0.039;
Matches 70; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 319 TTTGTAATCTAACAACAGATGCAATAGTTTAAATAAACTTATACTTAACTTTTAAAAAAA 378
|||| | | | |||| ||||| |||| ||| |||| || |||||| | | |||||
Db 125 TTTGAACTTTCACAAGAGATGTAATAATTTTGATAATAAAATTCTTAACCATAATAAAAA 184

Qy 379 AAAAAAAAAAAAAAAAAAAAAAAAAAAAA 406
|||||||||||||||||||||||||||||
Db 185 AAAAAAAAAAAAAAAAAAAAAAAAAAAAA 212

RESULT 4
ABV56953
ID ABV56953 standard; cDNA; 236 BP.
XX
AC ABV56953;
XX
DT 17-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 56944.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodyanamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US05171.
XX
PR 17-FEB-2000; 2000US-183319P.
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
DR WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer -
XX
PS Claim 1; Page 10969; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (I) is also useful as a pharmacodyanamic or pharmacogenomic marker.
XX
SQ Sequence 236 BP; 108 A; 33 C; 40 G; 53 T; 2 other;

Query Match 14.6%; Score 59.2; DB 23; Length 236;
Best Local Similarity 79.5%; Pred. No. 0.039;
Matches 70; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 319 TTTGTAATCTAACAACAGATGCAATAGTTTAAATAAACTTATACTTAACTTTTAAAAAAA 378
|||| | | | |||| ||||| |||| ||| |||| || |||||| | | |||||
Db 126 TTTGAACTTTCACAAGAGATGTAATAATTTTGATAATAAAATTCTTAACCATAATAAAAA 185

Qy 379 AAAAAAAAAAAAAAAAAAAAAAAAAAAAA 406
|||||||||||||||||||||||||||||
Db 186 AAAAAAAAAAAAAAAAAAAAAAAAAAAAA 213

RESULT 5
ABV57502
ID ABV57502 standard; cDNA; 574 BP.
XX
AC ABV57502;
XX
DT 17-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 57493.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodyanamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.

```
XX
PF   20-FEB-2001; 2001WO-US05171.
XX
PR   17-FEB-2000; 2000US-183319P.
PR   16-MAR-2000; 2000US-189862P.
PR   25-MAY-2000; 2000US-207454P.
PR   09-JUN-2000; 2000US-211314P.
PR   18-JUL-2000; 2000US-219007P.
PR   13-DEC-2000; 2000US-255281P.
XX
PA   (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI   Schlegel R,  Endege WO,  Monahan JE;
XX
DR   WPI; 2001-662795/76.
XX
PT   Novel isolated nucleic acid molecule associated with cancerous state of
PT   prostate cells and correlating with presence of prostate cancer, useful
PT   for detecting presence of prostate cancer, stage of prostate cancer  -
XX
PS   Claim 1; Page 11059; 11750pp; English.
XX
CC   The invention relates to an isolated nucleic acid molecule (I) comprising
CC   a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC   specification or its complement. (I) is useful for:
CC   (a) assessing whether a patient is afflicted with prostate cancer;
CC   (b) monitoring the progression of prostate cancer in a patient;
CC   (c) assessing the efficacy of a test compound to inhibit prostate
CC   cancer in a patient;
CC   (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC   in a patient;
CC   (e) selecting a composition for inhibiting prostate cancer in a patient;
CC   (f) assessing the prostate cell carcinogenic potential of a compound;
CC   (g) determining whether prostate cancer has metastasized in a patient;
CC   (h) assessing the aggressiveness or indolence of prostate cancer in a
CC   patient;
CC   (I) is also useful as a pharmacodyanamic or pharmacogenomic marker.
XX
SQ   Sequence 574 BP; 241 A; 113 C; 95 G; 124 T; 1 other;

  Query Match             14.3%;  Score 58;  DB 23;  Length 574;
  Best Local Similarity   58.8%;  Pred. No. 0.064;
  Matches  100;  Conservative    0;  Mismatches   70;  Indels    0;  Gaps    0;

Qy     237 ATTTATTTGCCATAATGGAGACCCGTTTTTATTGAATATCGTCAGTTTCCAATTAAAGTC 296
           |||| | |  || || ||  |    ||| |||  | |  |  |||  | |||| ||| |
Db     270 ATTTCTATTTCACAAGGGTAATTGTTTTATATACACTGGCAGCAGCATACAATAAAACTT 329

Qy     297 ATTTCGAGCCATACTGAATAATTTTGTAATCTAACAACAGATGCAATAGTTTAAATAAAC 356
           | |  ||  | |    || ||      ||   || || | |   || |    ||| |||
Db     330 AGTATGAAACTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 389

Qy     357 TTATACTTAACTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 406
             | |   ||     |||||||||||||||||||||||||||||||||
Db     390 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 439


RESULT 6
ABV58273
ID   ABV58273 standard; cDNA; 497 BP.
XX
AC   ABV58273;
XX
DT   13-SEP-2002  (first entry)
XX
DE   Human prostate expression marker cDNA 58264.
XX
KW   Human; prostate cancer; cytostatic; carcinogen; pharmacodyanamic marker;
KW   pharmacogenomic marker; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200160860-A2.
XX
PD   23-AUG-2001.
XX
PF   20-FEB-2001; 2001WO-US05171.
XX
PR   17-FEB-2000; 2000US-183319P.
PR   16-MAR-2000; 2000US-189862P.
PR   25-MAY-2000; 2000US-207454P.
PR   09-JUN-2000; 2000US-211314P.
PR   18-JUL-2000; 2000US-219007P.
PR   13-DEC-2000; 2000US-255281P.
XX
PA   (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI   Schlegel R,  Endege WO,  Monahan JE;
XX
DR   WPI; 2001-662795/76.
XX
PT   Novel isolated nucleic acid molecule associated with cancerous state of
PT   prostate cells and correlating with presence of prostate cancer, useful
PT   for detecting presence of prostate cancer, stage of prostate cancer  -
XX
PS   Claim 1; Page 11189; 11750pp; English.
XX
CC   The invention relates to an isolated nucleic acid molecule (I) comprising
CC   a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC   specification or its complement. (I) is useful for:
CC   (a) assessing whether a patient is afflicted with prostate cancer;
CC   (b) monitoring the progression of prostate cancer in a patient;
CC   (c) assessing the efficacy of a test compound to inhibit prostate
CC   cancer in a patient;
CC   (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC   in a patient;
CC   (e) selecting a composition for inhibiting prostate cancer in a patient;
CC   (f) assessing the prostate cell carcinogenic potential of a compound;
CC   (g) determining whether prostate cancer has metastasized in a patient;
CC   (h) assessing the aggressiveness or indolence of prostate cancer in a
CC   patient;
CC   (I) is also useful as a pharmacodyanamic or pharmacogenomic marker.
XX
SQ   Sequence 497 BP; 235 A; 77 C; 69 G; 115 T; 1 other;

  Query Match             14.0%;  Score 57;  DB 23;  Length 497;
  Best Local Similarity   56.8%;  Pred. No. 0.096;
  Matches  105;  Conservative    0;  Mismatches   80;  Indels    0;  Gaps    0;

Qy     222 ATAGGTGAAAATCCGATTTATTTGCCATAATGGAGACCCGTTTTTATTGAATATCGTCAG 281
           |||  |  |||     |  ||| ||   ||    | | | ||||||||| | |   || ||
Db     165 ATATATATAAAAAAAACCTATCTGTATTAGACAAAAACTGTTTTTATTTATTTCTGTAAG 224

Qy     282 TTTCCAATTAAAGTCATTTCGAGCCATACTGAATAATTTTGTAATCTAACAACAGATGCA 341
             |  ||||||||   || |        |    | ||       ||   || || | |   |
Db     225 ATATCAATTAAATAAATATTTTAGTGGAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAA 284

Qy     342 ATAGTTTAAATAAACTTATACTTAACTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 401
           | |    ||| |||   | |   ||     ||||||||||||||||||||||||||||
Db     285 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 344

Qy     402 AAAAA 406
           |||||
Db     345 AAAAA 349


RESULT 7
ABV58662
ID   ABV58662 standard; cDNA; 381 BP.
XX
AC   ABV58662;
XX
DT   13-SEP-2002  (first entry)
XX
DE   Human prostate expression marker cDNA 58653.
XX
```

KW Human; prostate cancer; cytostatic; carcinogen; pharmacodyanamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US05171.
XX
PR 17-FEB-2000; 2000US-183319P.
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
DR WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer -
XX
PS Claim 1; Page 11255; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (I) is also useful as a pharmacodyanamic or pharmacogenomic marker.
XX
SQ Sequence 381 BP; 196 A; 70 C; 49 G; 66 T; 0 other;

Query Match 13.8%; Score 56; DB 23; Length 381;
Best Local Similarity 77.3%; Pred. No. 0.15;
Matches 68; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Qy 319 TTTGTAATCTAACAACAGATGCAATAGTTTAAATAAACTTATACTTAACTTTTAAAAAAA 378
Db 112 TTTGAACTTTCACAAGAGATGTAATAATTTTGATAATAAAATTCTTAACCATAATAATCA 171

Qy 379 AAAAAAAAAAAAAAAAAAAAAAAAAAAAA 406
Db 172 AAAAAAAAAAAAAAAAAAAAAAAAAAAAA 199

RESULT 8
ABL32986/c
ID ABL32986 standard; DNA; 17294 BP.
XX
AC ABL32986;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human immune system associated gene SEQ ID NO: 959.
XX
KW Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antianaemic; cytostatic; nootropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200200928-A2.
XX
PD 03-JAN-2002.
XX
PF 02-JUL-2001; 2001WO-EP07537.
XX
PR 30-JUN-2000; 2000DE-1032529.
PR 01-SEP-2000; 2000DE-1043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-130909/17.
XX
PT Nucleic acid comprising fragment of chemically modified gene, useful
PT for diagnosis and treatment of diseases associated with abnormal
PT cytosine methylation -
XX
PS Claim 1; SEQ ID NO 959; 32pp + Sequence Listing; German.
XX
CC The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention.
XX
SQ Sequence 17294 BP; 4922 A; 203 C; 3691 G; 8478 T; 0 other;

Query Match 13.7%; Score 55.8; DB 24; Length 17294;
Best Local Similarity 53.4%; Pred. No. 0.16;
Matches 117; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

Qy 184 CTTTTGCAGAGGAACTTTCAGGACAACCTGCGTTTGCTATAGGTGAAAATCCGATTTATT 243
Db 280 CTTTAACCTATAAAAAAAAAAAATAAACCCCTATTAATTTAAATCACCATCAATTCGATT 221

Qy 244 TGCCATAATGGAGACCCGTTTTTATTGAATATCGTCAGTTTCCAATTAAAGTCATTTCGA 303
Db 220 TTCTATTAATCGTACAAATTTTTTTTTTTTTTTTTTTTTTTTTACTCAATATCAAACCA 161

Qy 304 GCCATACTGAATAATTTTGTAATCTAACAACAGATGCAATAGTTTAAATAAACTTATACT 363
Db 160 ACCTAAATAAATTAACTTTCAAAAAATTAAAATATAAACCATAATAAAAAAAAAAATAAC 101

Qy 364 TAACTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 402
Db 100 TAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAACAACAAA 62

RESULT 9
AAS29049
ID AAS29049 standard; cDNA; 325 BP.
XX
AC AAS29049;
XX
DT 21-NOV-2001 (first entry)
XX
DE cDNA encoding for human DNA-binding protein #20.
XX
KW Human; DNA-binding protein; histone; chromo domain protein;

KW chromatin organisation modifier; Y-box binding protein;
KW DNA organisation; gene transcription; malignant disease;
KW autoimmune disorder; rheumatic disease; genetic abnormality;
KW infectious disease; neurological disorder; gene therapy;
KW immunomodulatory; anti-HIV; anti rheumatic; anti microbial;
KW cytostatic; ss.
XX
OS Homo sapiens.
XX
PN WO200155162-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01305.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.

PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-465557/50.
DR P-PSDB; AAU18173.
XX
PT Nucleic acid molecules encoding human secreted chromosomal binding
PT proteins, used in preventing, treating or ameliorating a disorder, e.g.
PT Alzheimer's and Parkinson's diseases and cancers -
XX
PS Claim 4; SEQ ID No 30; 561pp; English.
XX
CC The present invention relates to the isolation of novel DNA-binding
CC proteins (AAU18154-AAU18281), and cDNA and genomic sequences encoding
CC for these proteins. DNA-binding proteins such as histones, chromo
CC (chromatin organisation modifier) domain proteins, and Y-box binding
CC proteins may contribute to diseases resulting from aberrant DNA
CC organisation and/or gene transcription. The sequences of the invention
CC are useful in screening assays to identify antagonists and/or agonists
CC that may enhance or block activities mediated by DNA-binding proteins.
CC Blockers of DNA-binding proteins may be useful in treating disorders
CC such as malignant diseases (e.g. cancer), autoimmune disorders
CC (e.g. diabetes mellitus), rheumatic diseases (e.g. rheumatoid
CC arthritis), genetic abnormalities (e.g. cystic fibrosis), infectious
CC diseases (e.g. HIV) and neurological disorders (e.g. Alzheimer's
CC disease). The polynucleotide sequences of the invention may also be
CC used in gene therapy. AAS29030-AAS29157 represent cDNA sequences
CC encoding for novel DNA-binding proteins.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 325 BP; 174 A; 38 C; 36 G; 77 T; 0 other;

Query Match 13.6%; Score 55.2; DB 22; Length 325;
Best Local Similarity 60.8%; Pred. No. 0.2;
Matches 90; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

```
Qy    259 CCGTTTTTATTGAATATCGTCAGTTTCCAATTAAAGTCATTTCGAGCCATACTGAATAAT 318
          || || | ||| | ||   | |     | |||| |||   |  ||  ||    ||| ||
Db    104 CCTTTATAATTCACTAAAATAAAGCATCTATTAGTGTCTGATTTAGGAATGTAAAATGAT 163

Qy    319 TTTGTAATCTAACAACAGATGCAATAGTTTAAATAAACTTATACTTAACTTTTAAAAAAA 378
          | |||| |     ||   |    ||   ||  | |||  |||    ||| |  |||||||
Db    164 TCTGTATTAATGTAAATAAGATTATCTATTGCAAAAAGATATTTCAAACCTAAAAAAAAA 223

Qy    379 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 406
          ||||||||||||||||||||||||||||||
Db    224 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 251
```

RESULT 10
AAS29124
ID AAS29124 standard; cDNA; 621 BP.
XX
AC AAS29124;
XX
DT 21-NOV-2001 (first entry)
XX
DE cDNA encoding for human DNA-binding protein #95.
XX
KW Human; DNA-binding protein; histone; chromo domain protein;
KW chromatin organisation modifier; Y-box binding protein;
KW DNA organisation; gene transcription; malignant disease;
KW autoimmune disorder; rheumatic disease; genetic abnormality;
KW infectious disease; neurological disorder; gene therapy;
KW immunomodulatory; anti-HIV; anti rheumatic; anti microbial;
KW cytostatic; ss.
XX
OS Homo sapiens.
XX
PN WO200155162-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01305.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.

PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-465557/50.
DR P-PSDB; AAU18248.
XX
PT Nucleic acid molecules encoding human secreted chromosomal binding
PT proteins, used in preventing, treating or ameliorating a disorder, e.g.
PT Alzheimer's and Parkinson's diseases and cancers -
XX
PS Claim 4; SEQ ID No 105; 561pp; English.
XX
CC The present invention relates to the isolation of novel DNA-binding
CC proteins (AAU18154-AAU18281), and cDNA and genomic sequences encoding
CC for these proteins. DNA-binding proteins such as histones, chromo
CC (chromatin organisation modifier) domain proteins, and Y-box binding
CC proteins may contribute to diseases resulting from aberrant DNA
CC organisation and/or gene transcription. The sequences of the invention
CC are useful in screening assays to identify antagonists and/or agonists
CC that may enhance or block activities mediated by DNA-binding proteins.
CC Blockers of DNA-binding proteins may be useful in treating disorders
CC such as malignant diseases (e.g. cancer), autoimmune disorders
CC (e.g. diabetes mellitus), rheumatic diseases (e.g. rheumatoid
CC arthritis), genetic abnormalities (e.g. cystic fibrosis), infectious
CC diseases (e.g. HIV) and neurological disorders (e.g. Alzheimer's
CC disease). The polynucleotide sequences of the invention may also be
CC used in gene therapy. AAS29030-AAS29157 represent cDNA sequences
CC encoding for novel DNA-binding proteins.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 621 BP; 250 A; 123 C; 96 G; 146 T; 6 other;

Query Match 13.6%; Score 55.2; DB 22; Length 621;
Best Local Similarity 60.8%; Pred. No. 0.2;
Matches 90; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

```
Qy    259 CCGTTTTTATTGAATATCGTCAGTTTCCAATTAAAGTCATTTCGAGCCATACTGAATAAT 318
          || || | ||| | ||   | |     | |||| |||   |  ||  ||   ||| ||
Db    388 CCTTTATAATTCACTAAAATAAAGCATCTATTAGTGTCTGATTTAGGAATGTAAAATGAT 447

Qy    319 TTTGTAATCTAACAACAGATGCAATAGTTTAAATAAACTTATACTTAACTTTTAAAAAAA 378
          | |||| |     ||   |    ||   ||  | |||  |||    ||| |  |||||||
Db    448 TCTGTATTAATGTAAATAAGATTATCTATTGCAAAAAGATATTTCAAACCTAAAAAAAAA 507

Qy    379 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 406
          ||||||||||||||||||||||||||||||
Db    508 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 535
```

RESULT 11
AAH77211
ID AAH77211 standard; cDNA; 351 BP.
XX
AC AAH77211;
XX
DT 30-JAN-2002 (first entry)
XX
DE Arthropod defensin cDNA from Argiope sp. clone aot1c.pk0 #2.
XX
KW Arthropod; defensin; aot1c.pk0; antimicrobial; microbial pathogen;
KW gene therapy; ss.
XX
OS Argiope sp.
XX
FH Key Location/Qualifiers
FT CDS 28..210
FT /*tag= a
FT /product= "Defensin"

XX
PN EP1146052-A2.
XX
PD 17-OCT-2001.
XX
PF 12-APR-2001; 2001EP-0303488.
XX
PR 14-APR-2000; 2000US-197279P.
XX
PA (DUPO) DU PONT DE NEMOURS & CO E I.
XX
PI Presnail JK, Weng Z, Wong JF;
XX
DR WPI; 2002-001056/01.
DR P-PSDB; AAG77934.
XX
PT New polynucleotides encoding Arthropod defensin polypeptides for
PT improving the microbial resistance of plants and animals by gene
PT therapy -
XX
PS Claim 6; Page 19; 22pp; English.
XX
CC The sequence represents a novel Arthropod defensin from Ariope sp.
CC clone aotlc.pk0. The invention relates to novel polynucleotides encoding
CC Arthropod defensin polypeptides. The polypeptides and polynucleotides of
CC the invention have antimicrobial activity. The polynucleotides may be
CC used to genetically engineer cells and organisms, especially plants, to
CC alter their resistance to a wide range of microbial pathogens. The
CC polynucleotides may also have a use in gene therapy.
XX
SQ Sequence 351 BP; 95 A; 84 C; 69 G; 103 T; 0 other;

Query Match 13.4%; Score 54.6; DB 24; Length 351;
Best Local Similarity 54.9%; Pred. No. 0.26;
Matches 130; Conservative 0; Mismatches 104; Indels 3; Gaps 1;

```
Qy      23 TTTCTCCACTCAGCTTCAAGAATGAAATCCATAGCTATTATTTTCATCGTTCTTGTTGCC 82
           | | || ||    |   || ||||||  | | || | |  | || |||   || ||   |
Db       7 TCTGTCGACATTTCCAAAAAAATGAATGCGAGAGTTCTGTTGTTGATCTGCCTAGT---C 63

Qy      83 TTCTGTATTTTGGAGGATGGGATTGTAGAAGCTGGTTTTGGATGTCCCTTTAATGCAGGA 142
            |||||  ||| |     | ||  || ||||||||||| || || |||||  |
Db      64 GTCTGTGCTTTTGCCACAGTGACCGTGGAAGCTGGTTTCGGCTGCCCCTTCGACCAGATG 123

Qy     143 AAATGCCATAGACATTGCAAAAGTATTCGTCGTAGAGGAGGCTTTTGCAGAGGAACTTTC 202
           | || || |  |||||||  || ||      ||| ||||| |  ||||         |||
Db     124 CAGTGTCACAATCATTGCAGGAGCATCAAATATAGGGGAGGATACTGCACCAACTTATTC 183

Qy     203 AGGACAACCTGCGTTTGCTATAGGTGAAAATCCGATTTATTTGCCATAATGGAGACC 259
           | |   ||||||   || ||  | |||  | ||   |    |  || |  |||| ||
Db     184 AAGCGCACCTGCAAGTGTTACGGATGATGACCCCCCTCCCCTCTCAGACAGGAGCCC 240
```

RESULT 12
AAS60450/c
ID AAS60450 standard; cDNA; 425 BP.
XX
AC AAS60450;
XX
DT 29-JAN-2002 (first entry)
XX
DE Human cancer agent-sensitive marker #181.
XX
KW Human; cancer cell marker; TAXOL; cytostatic; tumour; carcinoma;
KW squamous cell carcinoma; sarcoma; fibrosarcoma; leukaemia;
KW lymphocytic leukaemia; lymphoma; plasmocytoma; reticulum cell sarcoma;
KW Hodgkin's disease; glioma; ss.
XX
OS Homo sapiens.
XX
PN WO200179556-A2.
XX
PD 25-OCT-2001.
XX
PF 13-APR-2001; 2001WO-US12132.
XX
PR 14-APR-2000; 2000US-197538P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Lillie J, Brown JL, Bolt A, Van Huffel C;
XX
DR WPI; 2001-602933/68.
XX
PT Novel nucleic acid, used as a marker to determine the effectiveness of
PT using TAXOL to treat cancer cell growth in individuals -
XX
PS Claim 1; Page 221; 527pp; English.
XX
CC The invention relates to 1046 novel nucleic acids which are used as
CC markers for determining the sensitivity of a cancer cell to the
CC anticancer agent TAXOL. Cancer cells can be treated with TAXOL when
CC they are shown to express one of the 242 sensitivity markers or the
CC cells are shown not to express one of the 804 resistance markers.
CC The methods can be used to determine the effectiveness of TAXOL
CC in the treatment of cancer cell growth in an individual. The markers
CC can be used as targets in developing anti-cancer agents such as
CC chemotherapeutic compounds. The markers can also be used as targets in
CC developing treatments for cancer, particularly those cancers which
CC display resistance to agents and exhibit expression of the markers. The
CC anticancer agents developed by the novel method can be used to treat
CC cancer. Probes based on the markers can be used to detect transcripts or
CC genomic sequences corresponding to the markers, in the identification of
CC cells or tissues which mis-express the protein. Cancers which may
CC be targeted include carcinoma (e.g. squamous cell carcinoma),
CC sarcoma (e.g. fibrosarcoma) leukaemia (e.g. lymphocytic leukaemia),
CC lymphoma, plasmocytoma, reticulum cell sarcoma, Hodgkin's disease and
CC tumours (e.g. glioma). The present sequence is one of the 1046
CC novel cancer cell markers.
XX
SQ Sequence 425 BP; 159 A; 41 C; 28 G; 194 T; 3 other;

Query Match 13.4%; Score 54.6; DB 22; Length 425;
Best Local Similarity 53.4%; Pred. No. 0.26;
Matches 111; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

```
Qy     199 TTTCAGGACAACCTGCGTTTGCTATAGGTGAAAATCCGATTTATTTGCCATAATGGAGAC 258
           ||| |  | || | | |   |  | |    |||| ||  ||| |||    |   || ||
Db     215 TTTTAAAAAAAACGGGGGGGGGAAAAAAAAAAAAACCCTTTTTTTTTTTTTTTGGGGGAA 156

Qy     259 CCGTTTTTATTGAATATCGTCAGTTTCCAATTAAAGTCATTTCGAGCCATACTGAATAAT 318
           |||||  ||   | | |   ||   | ||    |  |||      | |   || ||
Db     155 ATTTTTTTTTTTTTTTTTTTTTTTAAAATTTTTTTTTTTTTTNNNAAAAAAAAAAAAAA 96

Qy     319 TTTGTAATCTAACAACAGATGCAATAGTTTAAATAAACTTATACTTAACTTTTAAAAAAA 378
              ||   || || | |   || |    ||| |||   | |   ||     |||||||
Db      95 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 36

Qy     379 AAAAAAAAAAAAAAAAAAAAAAAAAA 406
           ||||||||||||||||||||||||||
Db      35 AAAAAAAAAAAAAAAAAAAAAAAAAA 8
```

RESULT 13
ABQ54618
ID ABQ54618 standard; cDNA; 2736 BP.
XX
AC ABQ54618;
XX
DT 22-AUG-2002 (first entry)
XX
DE Human ovarian antigen HDPPN96 cDNA, SEQ ID NO:498.
XX
KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
KW ovarian cancer; breast cancer; tumour; reproductive system disorder;
KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;

```
KW   PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
KW   inflammatory condition; immune disorder; blood disorder;
KW   cardiovascular disorder; respiratory disorder; neurological disorder;
KW   gastrointestinal disorder; urinary system disorder; drug screening;
KW   gene therapy; chromosome mapping; forensic analysis;
KW   antibody preparation; cytostatic; immunomodulatory; neuroprotective;
KW   antiinflammatory; gynaecological; reproductive; chromosome 1q22;
KW   gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200200677-A1.
XX
PD   03-JAN-2002.
XX
PF   07-JUN-2001; 2001WO-US18569.
XX
PR   07-JUN-2000; 2000US-209467P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Birse CE,  Rosen CA;
XX
DR   WPI; 2002-147878/19.
DR   P-PSDB; ABP41541.
XX
PT   Isolated nucleic acid molecules encoding novel ovarian polypeptides,
PT   useful in the prevention, treatment and diagnosis of cancer (e.g.
PT   ovarian cancer), immune disorders, cardiovascular disorders and
PT   neurological diseases -
XX
PS   Claim 1; SEQ ID No 498; 2922pp; English.
XX
CC   The invention relates to 2175 novel human ovarian antigens (ABP41054-
CC   ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
CC   encompasses polypeptides 90% identical and polynucleotides 95% identical
CC   to the sequences of the invention. The invention additionally relates to
CC   recombinant vectors and host cells comprising human ovarian antigen
CC   polynucleotides, antibodies against human ovarian antigens, and the use
CC   of ovarian antigen polynucleotides and polypeptides in diagnosing,
CC   treating, prognosing or preventing various ovary and/or breast-related
CC   disorders. Such conditions include ovarian cancer and breast cancer, and
CC   metastatic tumours of ovarian or breast origin, reproductive system
CC   disorders (e.g., infertility, disorders of pregnancy, anovulation,
CC   polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
CC   disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
CC   shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
CC   vaginitis), immune disorders (e.g., congenital and acquired
CC   immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
CC   blood-related disorders (e.g., anaemia), cardiovascular disorders,
CC   respiratory disorders, neurological disorders, gastrointestinal disorders
CC   and urinary system disorders. Ovarian antigen polypeptides and
CC   polynucleotides may also be used in screening for compounds which
CC   modulate ovarian antigen expression or activity. The polynucleotides may
CC   further be used for gene therapy, chromosome mapping, in the
CC   identification of individuals and in forensic analysis, and the
CC   polypeptides may be used as food additives or to prepare antibodies
CC   useful in disease diagnosis, drug targeting and phenotyping. The present
CC   sequence represents cDNA encoding a human ovarian antigen of the
CC   invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 2736 BP; 989 A; 528 C; 543 G; 672 T; 4 other;

  Query Match             13.4%;  Score 54.6;  DB 24;  Length 2736;
  Best Local Similarity   48.7%;  Pred. No. 0.26;
  Matches  147;  Conservative    0;  Mismatches  155;  Indels    0;  Gaps    0;

Qy     105 TTGTAGAAGCTGGTTTTGGATGTCCCTTTAATGCAGGAAAATGCCATAGACATTGCAAAA 164
           || || || |||| | |   || |  |    |  | | | ||    |  |  |  | |||
Db    2432 TTCTAAAATCTGGATGTCATTGACGATAATGTTTATGGAGATAAGGTCTAAGTGCCTAAA 2491

Qy     165 GTATTCGTCGTAGAGGAGGCTTTTGCAGAGGAACTTTCAGGACAACCTGCGTTTGCTATA 224
           | |   | || |   || |     | |  |   | ||     |||  | | |    || .
Db    2492 AAAATGTACATATACCTGGTTGAAATACAACACTATACATACACACCACCATATATACTA 2551

Qy     225 GGTGAAAATCCGATTTATTTGCCATAATGGAGACCCGTTTTTATTGAATATCGTCAGTTT 284
           | ||  ||||| ||  | | |       ||||  |  |||| |||    || ;||   |||
Db    2552 GCTGTTAATCCTATGGAATGGGGTATTGGGAGTGCTTTTTTAATTTTTCATAGTTTTTTT 2611

Qy     285 CCAATTAAAGTCATTTCGAGCCATACTGAATAATTTTGTAATCTAACAACAGATGCAATA 344
           ||| |||     |      |||    ||    | |    |  || || | ||  |
Db    2612 TTAATAAAATGGCATATTTTGCATCTACAACTTCTATAATTTGAAAAAATAAATAAACAT 2671

Qy     345 GTTTAAATAAACTTATACTTAACTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 404
           |    |      | |   ||     ||||||||||||||||||||| ||||||||||
Db    2672 TATCTTTTTTGTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANAAAAAAAAAA 2731

Qy     405 AA 406
           ||
Db    2732 AA 2733


RESULT 14
ABV37528/c
ID   ABV37528 standard; cDNA; 372 BP.
XX
AC   ABV37528;
XX
DT   16-SEP-2002  (first entry)
XX
DE   Human prostate expression marker cDNA 37519.
XX
KW   Human; prostate cancer; cytostatic; carcinogen; pharmacodyanamic marker;
KW   pharmacogenomic marker; gene; ss.
XX
OS   Homo sapiens.
XX
PN   WO200160860-A2.
XX
PD   23-AUG-2001.
XX
PF   20-FEB-2001; 2001WO-US05171.
XX
PR   17-FEB-2000; 2000US-183319P.
PR   16-MAR-2000; 2000US-189862P.
PR   25-MAY-2000; 2000US-207454P.
PR   09-JUN-2000; 2000US-211314P.
PR   18-JUL-2000; 2000US-219007P.
PR   13-DEC-2000; 2000US-255281P.
XX
PA   (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI   Schlegel R,  Endege WO,  Monahan JE;
XX
DR   WPI; 2001-662795/76.
XX
PT   Novel isolated nucleic acid molecule associated with cancerous state of
PT   prostate cells and correlating with presence of prostate cancer, useful
PT   for detecting presence of prostate cancer, stage of prostate cancer  -
XX
PS   Claim 1; Page 7697-7698; 11750pp; English.
XX
CC   The invention relates to an isolated nucleic acid molecule (I) comprising
CC   a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC   specification or its complement. (I) is useful for:
CC   (a) assessing whether a patient is afflicted with prostate cancer;
CC   (b) monitoring the progression of prostate cancer in a patient;
CC   (c) assessing the efficacy of a test compound to inhibit prostate
CC   cancer in a patient;
CC   (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC   in a patient;
CC   (e) selecting a composition for inhibiting prostate cancer in a patient;
CC   (f) assessing the prostate cell carcinogenic potential of a compound;
```

CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (I) is also useful as a pharmacodyanamic or pharmacogenomic marker.
XX
SQ Sequence 372 BP; 141 A; 37 C; 50 G; 144 T; 0 other;

Query Match 13.4%; Score 54.4; DB 23; Length 372;
Best Local Similarity 57.7%; Pred. No. 0.28;
Matches 97; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

Qy 239 TTATTTGCCATAATGGAGACCCGTTTTTATTGAATATCGTCAGTTTCCAATTAAAGTCAT 298
Db 295 TTTTTTTTTATTTTAAAATTACATTTTTTAATTCTTTTTAAAGCCCATTATTATTTTTTC 236

Qy 299 TTCGAGCCATACTGAATAATTTTGTAATCTAACAACAGATGCAATAGTTTAAATAAACTT 358
Db 235 TTTTAATAATTTTTTTTTTTTTTTTAATATACCCCTTTTTTTTTTAATTTAAATAAATTT 176

Qy 359 ATACTTAACTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 406
Db 175 TTTTTTTGGTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 128

RESULT 15
ABV10351/c
ID ABV10351 standard; cDNA; 447 BP.
XX
AC ABV10351;
XX
DT 13-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 10342.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodyanamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US05171.
XX
PR 17-FEB-2000; 2000US-183319P.
PR 16-MAR-2000; 2000US-189862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
DR WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer -
XX
PS Claim 1; Page 1663; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (I) is also useful as a pharmacodyanamic or pharmacogenomic marker.
XX
SQ Sequence 447 BP; 98 A; 75 C; 91 G; 183 T; 0 other;

Query Match 13.4%; Score 54.4; DB 23; Length 447;
Best Local Similarity 65.8%; Pred. No. 0.28;
Matches 79; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

Qy 287 AATTAAAGTCATTTCGAGCCATACTGAATAATTTTGTAATCTAACAACAGATGCAATAGT 346
Db 428 AATAAAATAAAATTAAAATAAAACTTAAAATATAATCAAATTAATATATAAAATCAAAAA 369

Qy 347 TTAAATAAACTTATACTTAACTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 406
Db 368 TAAAATAAACTAAATTTAAACTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 309

Search completed: May 5, 2003, 19:18:44
Job time : 238 secs

GenCore version 5.1.5

OM nucleic - nucleic search, using sw model

Run on: May 5, 2003, 19:18:48 ; Search time 50 Seconds
(without alignments)
2490.215 Million cell updates/sec

Title: US-09-829-481-3
Perfect score: 406
Sequence: 1 ctctactacaatcactaagt.....aaaaaaaaaaaaaaaaaaaa 406

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	55.6	13.7	1652	4	US-09-627-650B-2	Sequence 2, Appli
2	55.6	13.7	1652	4	US-09-436-063C-2	Sequence 2, Appli
3	52.8	13.0	3581	2	US-08-738-349-1	Sequence 1, Appli
4	52.4	12.9	958	2	US-08-757-046A-5	Sequence 5, Appli
5	52.4	12.9	958	3	US-09-447-208-5	Sequence 5, Appli
6	52.4	12.9	958	3	US-09-135-988-5	Sequence 5, Appli
7	52.4	12.9	958	4	US-09-277-716-5	Sequence 5, Appli
8	52.4	12.9	958	4	US-08-597-274A-5	Sequence 5, Appli
9	52.4	12.9	958	4	US-08-908-909-5	Sequence 5, Appli
10	52.4	12.9	958	4	US-09-609-161B-5	Sequence 5, Appli
11	52.4	12.9	958	4	US-08-990-103-5	Sequence 5, Appli
12	52	12.8	857	1	US-08-308-883-1	Sequence 1, Appli
13	52	12.8	857	1	US-08-730-163-1	Sequence 1, Appli
14	52	12.8	857	4	US-08-256-799-1	Sequence 1, Appli
15	52	12.8	857	4	US-08-462-437-1	Sequence 1, Appli
16	50.8	12.5	1582	3	US-08-545-196B-10	Sequence 10, Appl
17	50.8	12.5	1582	3	US-08-545-196B-12	Sequence 12, Appl
18	50.6	12.5	1411	4	US-08-964-127-5	Sequence 5, Appli
19	50.6	12.5	1411	4	US-09-496-692-5	Sequence 5, Appli
20	50.6	12.5	1641	1	US-08-300-903A-8	Sequence 8, Appli
21	50.4	12.4	2360	3	US-08-836-567-9	Sequence 9, Appli
22	50.4	12.4	3927	4	US-09-293-238B-1	Sequence 1, Appli
23	50.4	12.4	3933	1	US-08-199-776-1	Sequence 1, Appli
24	50.4	12.4	3933	3	US-08-663-731-1	Sequence 1, Appli
25	50.4	12.4	3933	3	US-08-879-338-1	Sequence 1, Appli
26	50.4	12.4	3933	5	PCT-US95-02044-1	Sequence 1, Appli
27	50.2	12.4	1474	4	US-08-821-994-64	Sequence 64, Appl
28	50.2	12.4	2790	3	US-08-800-291B-1	Sequence 1, Appli
29	50	12.3	1522	4	US-09-413-574-1	Sequence 1, Appli
30	49.6	12.2	2671	6	5168051-9	Patent No. 5168051
31	49.4	12.2	458	1	US-08-524-757-1	Sequence 1, Appli
32	48.4	11.9	841	4	US-09-004-731-40	Sequence 40, Appl
c 33	48.4	11.9	841	4	US-09-004-731-42	Sequence 42, Appl
34	48.4	11.9	841	4	US-09-032-215-46	Sequence 46, Appl
c 35	48.4	11.9	841	4	US-09-032-215-48	Sequence 48, Appl
36	48.4	11.9	841	4	US-08-749-699-40	Sequence 40, Appl
c 37	48.4	11.9	841	4	US-08-749-699-42	Sequence 42, Appl
38	48.4	11.9	841	4	US-09-004-729-40	Sequence 40, Appl
c 39	48.4	11.9	841	4	US-09-004-729-42	Sequence 42, Appl
40	48.4	11.9	1558	1	US-08-455-550-7	Sequence 7, Appli
41	48.4	11.9	1736	3	US-09-182-816-22	Sequence 22, Appl
c 42	48.4	11.9	1736	3	US-09-182-816-24	Sequence 24, Appl
43	48.4	11.9	1736	3	US-09-471-528-22	Sequence 22, Appl
c 44	48.4	11.9	1736	3	US-09-471-528-24	Sequence 24, Appl
45	48.4	11.9	1736	4	US-09-634-530-22	Sequence 22, Appl

ALIGNMENTS

RESULT 1
US-09-627-650B-2
; Sequence 2, Application US/09627650B
; Patent No. 6406872
; GENERAL INFORMATION:
; APPLICANT: Bamber, Bruce
; APPLICANT: Jorgensen, Erik
; TITLE OF INVENTION: Nematode Neuromuscular Junction GABA Receptors and
; TITLE OF INVENTION: Methods Related Thereto
; FILE REFERENCE: 21101.0009U3
; CURRENT APPLICATION NUMBER: US/09/627,650B
; CURRENT FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 09/436,063
; PRIOR FILING DATE: 1999-11-08
; PRIOR APPLICATION NUMBER: 60/107,727
; PRIOR FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1652
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
US-09-627-650B-2

Query Match 13.7%; Score 55.6; DB 4; Length 1652;
Best Local Similarity 55.1%; Pred. No. 0.0017;
Matches 109; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

```
Qy     209 ACCTGCGTTTGCTATAGGTGAAAATCCGATTTATTTGCCATAATGGAGACCCGTTTTTAT 268
           |||| | ||      |    || |||   | |||| ||||   | | | | ||||| |
Db    1450 ACCTTCATTAATCTCAATCCAACTTCCTCATCATTTTCCATTTCGAATATCTCTTTTTCT 1509

Qy     269 TGAATATCGTCAGTTTCCAATTAAAGTCATTTCGAGCCATACTGAATAATTTTGTAATCT 328
           || | |    | |||    ||    |   ||       | |    | ||||  |   |
Db    1510 TGCACAGAAGCCTTTTTTCGTTTTTTTTTATTGATTTATTTTTACGGATTTTTAGATAAT 1569

Qy     329 AACAACAGATGCAATAGTTTAAATAAACTTATACTTAACTTTTAAAAAAAAAAAAAAAAA 388
             | |      || | | ||||||| ||||  | |   |  ||||||||||||||||||
Db    1570 GCACAGATGCCTCATTGCTCAAATAAATTTATTTTAATTGTCGAAAAAAAAAAAAAAAAA 1629

Qy     389 AAAAAAAAAAAAAAAAA 406
           |||||||||||||||||
Db    1630 AAAAAAAAAAAAAAAAA 1647
```

RESULT 2
US-09-436-063C-2
; Sequence 2, Application US/09436063C
; Patent No. 6407210
; GENERAL INFORMATION:

```
;  APPLICANT: Bamber, Bruce
;  APPLICANT:  Jorgensen, Erik
;  TITLE OF INVENTION: Nematode Neuromuscular Junction GABA Receptors and
;  TITLE OF INVENTION:  Methods Related Thereto
;  FILE REFERENCE: P-1095corrected
;  CURRENT APPLICATION NUMBER: US/09/436,063C
;  CURRENT FILING DATE:  1999-11-08
;  PRIOR APPLICATION NUMBER: 60/107727
;  PRIOR FILING DATE: 1998-11-09
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;    LENGTH: 1652
;    TYPE: DNA
;    ORGANISM: Caenorhabditis elegans
US-09-436-063C-2

  Query Match             13.7%;  Score 55.6;  DB 4;  Length 1652;
  Best Local Similarity   55.1%;  Pred. No. 0.0017;
  Matches  109;  Conservative    0;  Mismatches   89;  Indels    0;  Gaps    0;

Qy     209 ACCTGCGTTTGCTATAGGTGAAAATCCGATTTATTTGCCATAATGGAGACCCGTTTTTAT 268
           |||| | ||       |    ||  |||   | |||| ||||   | | | |  ||||| |
Db    1450 ACCTTCATTAATCTCAATCCAACTTCCTCATCATTTTCCATTTCGAATATCTCTTTTTCT 1509

Qy     269 TGAATATCGTCAGTTTCCAATTAAAGTCATTTCGAGCCATACTGAATAATTTTGTAATCT 328
           || | |    | |||    ||    |  ||       | |    | ||||  |   |
Db    1510 TGCACAGAAGCCTTTTTTCGTTTTTTTTTATTGATTTATTTTTACGGATTTTTAGATAAT 1569

Qy     329 AACAACAGATGCAATAGTTTAAATAAACTTATACTTAACTTTTAAAAAAAAAAAAAAAAA 388
             | |       || | | ||||||| ||||  | |   |  ||||||||||||||||
Db    1570 GCACAGATGCCTCATTGCTCAAATAAATTTATTTTAATTGTCGAAAAAAAAAAAAAAAAA 1629

Qy     389 AAAAAAAAAAAAAAAAA 406
           |||||||||||||||||
Db    1630 AAAAAAAAAAAAAAAAA 1647


RESULT 3
US-08-738-349-1
; Sequence 1, Application US/08738349
; Patent No. 5869638
;  GENERAL INFORMATION:
;    APPLICANT:  Takeshita, Sunao
;    APPLICANT:  Okazaki, Makoto
;    APPLICANT:  Kawai, Shinji
;    APPLICANT:  Tsujimura, Atsushi
;    APPLICANT:  Amann, Egon
;    TITLE OF INVENTION:  Bone-Related Cadherin-Like Protein and
;    TITLE OF INVENTION:  Process for Its Production
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Finnegan, Henderson, Farabow, Garrett &
;      ADDRESSEE:  Dunner
;      STREET:  1300 I Street, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20005-3315
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/738,349
;      FILING DATE:  25-OCT-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/364,439
;      FILING DATE:
;      APPLICATION NUMBER:  US 08/112,061
;      FILING DATE:  26-AUG-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Barker, M. P.
;      REGISTRATION NUMBER:  32,013
;      REFERENCE/DOCKET NUMBER:  02481.1323-00000
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-408-4000
;      TELEFAX:  202-408-4400
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  3581 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA to mRNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Mus musculus
;      STRAIN:  osteoblastic cell line MC3T3E1
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  284..2671
US-08-738-349-1

  Query Match             13.0%;  Score 52.8;  DB 2;  Length 3581;
  Best Local Similarity   71.9%;  Pred. No. 0.0071;
  Matches   69;  Conservative    0;  Mismatches   27;  Indels    0;  Gaps    0;

Qy     311 TGAATAATTTTGTAATCTAACAACAGATGCAATAGTTTAAATAAACTTATACTTAACTTT 370
           ||  | | | ||   |  || |  ||| ||||   || |||||||   | ||   |||
Db    3443 TGCTTTAATATGAGCTTCAATATAAGAAGCAACCTTTGAAATAAAAAAAAGATTCTTTTT 3502

Qy     371 TAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 406
           |||||||||||||||||||||||||||||||||||||
Db    3503 TAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3538


RESULT 4
US-08-757-046A-5
; Sequence 5, Application US/08757046A
; Patent No. 5876995
;  GENERAL INFORMATION:
;    APPLICANT:  Bryan, Bruce
;    TITLE OF INVENTION:  BIOLUMINESCENT ARTICLES OF  MANUFACTURE
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brown, Martin, Haller & McClain
;      STREET:  1660 Union Street
;      CITY:  San Diego
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92101-2926
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/757,046A
;      FILING DATE:  11-25-96
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/597,274
;      FILING DATE:  02-06-96
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seidman, Stephanie L
;      REGISTRATION NUMBER:  33,779
;      REFERENCE/DOCKET NUMBER:  6680-105B
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-238-0999
;      TELEFAX:  619-238-0062
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  958 base pairs
```

```
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:
;    ORIGINAL SOURCE:
;    FEATURE:
;      NAME/KEY:  Coding Sequence
;      LOCATION:  115...702
;      OTHER INFORMATION:  apoaequorin-encoding gene
;    PUBLICATION INFORMATION:
;      AUTHORS:  Inouye et al.
;      JOURNAL:  Proc. Natl. Acad. Sci. U.S.A.
;      VOLUME:  82
;      PAGES:  3154-3158
;      DATE:  (1985)
;      DOCUMENT NUMBER:  PATENT NO.:  5,093,240
US-08-757-046A-5

  Query Match             12.9%;  Score 52.4;  DB 2;  Length 958;
  Best Local Similarity   57.2%;  Pred. No. 0.0072;
  Matches   95;  Conservative    0;  Mismatches   71;  Indels    0;  Gaps    0;

Qy    241 ATTTGCCATAATGGAGACCCGTTTTTATTGAATATCGTCAGTTTCCAATTAAAGTCATTT 300
          |||| ||| | |   || |  |||  || |  | |  | | |||   | | | |
Db    779 ATTTTCCAAATTTTTGAACGATTTCAATCGTTTGTGTTGATTTTTGTAATTAGGAACAGA 838

Qy    301 CGAGCCATACTGAATAATTTTGTAATCTAACAACAGATGCAATAGTTTAAATAAACTTAT 360
            |     | ||| || || | |   |  | |||||||     | |  |  || ||    |
Db    839 TTAAATCGAATGATTAGTTGTTTTTTTAATCAACAGAACTTACAAATCGAAAAAGTAAAA 898

Qy    361 ACTTAACTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 406
          |   ||     ||||||||||||||||||||||||||||||||||||||
Db    899 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 944


RESULT 5
US-09-447-208-5
; Sequence 5, Application US/09447208
; Patent No. 6113886
;  GENERAL INFORMATION:
;    APPLICANT:  Bryan, Bruce
;    TITLE OF INVENTION:  BIOLUMINESCENT ARTICLES OF  MANUFACTURE
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Heller Ehrman White & McAuliffe
;      STREET:  4250 Executive Square, 7th Floor
;      CITY:  La Jolla
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/447,208
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  0909/135,988
;      FILING DATE:  08-17-98
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/757,046
;      FILING DATE:  11-25-96
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/597,274
;      FILING DATE:  02-06-96
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seidman, Stephanie L
;      REGISTRATION NUMBER:  33,779
;      REFERENCE/DOCKET NUMBER:  24727-105C
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-450-8400
;      TELEFAX:  619-450-8499
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  958 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:
;    ORIGINAL SOURCE:
;    FEATURE:
;      NAME/KEY:  Coding Sequence
;      LOCATION:  115...702
;      OTHER INFORMATION:  apoaequorin-encoding gene
;    PUBLICATION INFORMATION:
;    PUBLICATION INFORMATION:  PATENT NO.:  5,093,240
;      AUTHORS:  Inouye et al.
;      JOURNAL:  Proc. Natl. Acad. Sci. U.S.A.
;      VOLUME:  82
;      PAGES:  3154-3158
;      DATE:  (1985)
US-09-447-208-5

  Query Match             12.9%;  Score 52.4;  DB 3;  Length 958;
  Best Local Similarity   57.2%;  Pred. No. 0.0072;
  Matches   95;  Conservative    0;  Mismatches   71;  Indels    0;  Gaps    0;

Qy    241 ATTTGCCATAATGGAGACCCGTTTTTATTGAATATCGTCAGTTTCCAATTAAAGTCATTT 300
          |||| ||| | |   || |  |||  || |  | |  | | |||   | | | |
Db    779 ATTTTCCAAATTTTTGAACGATTTCAATCGTTTGTGTTGATTTTTGTAATTAGGAACAGA 838

Qy    301 CGAGCCATACTGAATAATTTTGTAATCTAACAACAGATGCAATAGTTTAAATAAACTTAT 360
            |     | ||| || || | |   |  | |||||||     | |  |  || ||    |
Db    839 TTAAATCGAATGATTAGTTGTTTTTTTAATCAACAGAACTTACAAATCGAAAAAGTAAAA 898

Qy    361 ACTTAACTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 406
          |   ||     ||||||||||||||||||||||||||||||||||||||
Db    899 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 944


RESULT 6
US-09-135-988-5
; Sequence 5, Application US/09135988
; Patent No. 6152358
;  GENERAL INFORMATION:
;    APPLICANT:  Bryan, Bruce
;    TITLE OF INVENTION:  BIOLUMINESCENT ARTICLES OF  MANUFACTURE
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Heller Ehrman White & McAuliffe
;      STREET:  4250 Executive Square, 7th Floor
;      CITY:  La Jolla
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/135,988
;      FILING DATE:
;      CLASSIFICATION:
```

```
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/757,046
;      FILING DATE:  11-25-96
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/597,274
;      FILING DATE:  02-06-96
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seidman, Stephanie L
;      REGISTRATION NUMBER:  33,779
;      REFERENCE/DOCKET NUMBER:  24727-105C
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-450-8400
;      TELEFAX:  619-450-8499
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  958 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:
;    ORIGINAL SOURCE:
;    FEATURE:
;      NAME/KEY:  Coding Sequence
;      LOCATION:  115...702
;      OTHER INFORMATION:  apoaequorin-encoding gene
;    PUBLICATION INFORMATION:
;    PUBLICATION INFORMATION:  PATENT NO.:  5,093,240
;      AUTHORS:  Inouye et al.
;      JOURNAL:  Proc. Natl. Acad. Sci. U.S.A.
;      VOLUME:  82
;      PAGES:  3154-3158
;      DATE:  (1985)
US-09-135-988-5

  Query Match             12.9%;  Score 52.4;  DB 3;  Length 958;
  Best Local Similarity   57.2%;  Pred. No. 0.0072;
  Matches   95;  Conservative    0;  Mismatches   71;  Indels    0;  Gaps    0;

Qy     241 ATTTGCCATAATGGAGACCCGTTTTTATTGAATATCGTCAGTTTCCAATTAAAGTCATTT 300
           |||| ||| | |   || |  |||  || |  | |  | | |||   | | | |
Db     779 ATTTTCCAAATTTTTGAACGATTTCAATCGTTTGTGTTGATTTTTGTAATTAGGAACAGA 838

Qy     301 CGAGCCATACTGAATAATTTTTGTAATCTAACAACAGATGCAATAGTTTAAATAAACTTAT 360
             |     | ||| || || | |  |  | |||||||    | |  |  || ||    |
Db     839 TTAAATCGAATGATTAGTTGTTTTTTTAATCAACAGAACTTACAAATCGAAAAAGTAAAA 898

Qy     361 ACTTAACTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 406
           |   ||     |||||||||||||||||||||||||||||||||||
Db     899 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 944


RESULT 7
US-09-277-716-5
; Sequence 5, Application US/09277716A
; Patent No. 6232107
; GENERAL INFORMATION:
;  APPLICANT: Bryan, Bruce
;  APPLICANT: Szent-Gyorgyi, Christopher
;  APPLICANT: PROLUME, LTD.
;  TITLE OF INVENTION: LUCIFERASES, FLUORESCENT PROTEINS, NUCLEIC ACIDS ENCODING THE  I
;  CURRENT APPLICATION NUMBER: US/09/277,716A
;  CURRENT FILING DATE: 1999-03-26
;  EARLIER APPLICATION NUMBER: 60/102,939
;  EARLIER FILING DATE: 1998-10-01
;  EARLIER APPLICATION NUMBER: 60/089,367
;  EARLIER FILING DATE: 1998-06-15
;  EARLIER APPLICATION NUMBER: 60/079,624
;  EARLIER FILING DATE: 1998-03-27
;  NUMBER OF SEQ ID NOS: 32
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
;   LENGTH: 958
;   TYPE: DNA
;   ORGANISM: Aequorea (luminescent jellyfish)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (115)..(702)
;   FEATURE:
;   OTHER INFORMATION: Apoaequorin-encoding gene
;  PUBLICATION INFORMATION:
;   PATENT DOCUMENT NUMBER: 5,093,240
;   PATENT FILING DATE: 1987-10-08
;   PUBLICATION DATE: 1992-03-03
;  PUBLICATION INFORMATION:
;   AUTHORS: Inouye, S.
;   TITLE: Cloning and sequence analysis of cDNA for the luminescent protein    aequo
;   JOURNAL: Proc. Natl. Acad. Sci. USA
;   VOLUME: 82(10)
;   PAGES: 3154-3158
;   DATE: 1985-05
US-09-277-716-5

  Query Match             12.9%;  Score 52.4;  DB 4;  Length 958;
  Best Local Similarity   57.2%;  Pred. No. 0.0072;
  Matches   95;  Conservative    0;  Mismatches   71;  Indels    0;  Gaps    0;

Qy     241 ATTTGCCATAATGGAGACCCGTTTTTATTGAATATCGTCAGTTTCCAATTAAAGTCATTT 300
           |||| ||| | |   || |  |||  || |  | |  | | |||   | | | |
Db     779 ATTTTCCAAATTTTTGAACGATTTCAATCGTTTGTGTTGATTTTTGTAATTAGGAACAGA 838

Qy     301 CGAGCCATACTGAATAATTTTTGTAATCTAACAACAGATGCAATAGTTTAAATAAACTTAT 360
             |     | ||| || || | |  |  | |||||||    | |  |  || ||    |
Db     839 TTAAATCGAATGATTAGTTGTTTTTTTAATCAACAGAACTTACAAATCGAAAAAGTAAAA 898

Qy     361 ACTTAACTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 406
           |   ||     |||||||||||||||||||||||||||||||||||
Db     899 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 944


RESULT 8
US-08-597-274A-5
; Sequence 5, Application US/08597274A
; Patent No. 6247995
;  GENERAL INFORMATION:
;    APPLICANT:  Bryan, Bruce
;    TITLE OF INVENTION:  BIOLUMINESCENT NOVELTY ITEMS
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brown, Martin, Haller & McClain
;      STREET:  1660 Union Street
;      CITY:  San Diego
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92101-2926
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/597,274A
;      FILING DATE:  02/06/96
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seidman, Stephanie L
;      REGISTRATION NUMBER:  33,779
;      REFERENCE/DOCKET NUMBER:  6680-105
;    TELECOMMUNICATION INFORMATION:
```

```
;      TELEPHONE:  619-238-0999
;      TELEFAX:  619-238-0062
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  958 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:
;    ORIGINAL SOURCE:
;    FEATURE:
;      NAME/KEY:  Coding Sequence
;      LOCATION:  115...702
;      OTHER INFORMATION:  apoaequorin-encoding gene
;    PUBLICATION INFORMATION:
;      DOCUMENT NUMBER:  5,093,240
;      AUTHORS:  Inouye et al.
;      JOURNAL:  Proc. Natl. Acad. Sci. U.S.A.
;      VOLUME:  82
;      PAGES:  3154-3158
;      DATE:  (1985)
US-08-597-274A-5

  Query Match             12.9%;  Score 52.4;  DB 4;  Length 958;
  Best Local Similarity   57.2%;  Pred. No. 0.0072;
  Matches   95;  Conservative    0;  Mismatches   71;  Indels    0;  Gaps    0;

Qy     241 ATTTGCCATAATGGAGACCCGTTTTTATTGAATATCGTCAGTTTCCAATTAAAGTCATTT 300
           |||| ||| | |   || |  |||  || |  | |  | | |||   | | | |
Db     779 ATTTTCCAAATTTTTGAACGATTTCAATCGTTTGTGTTGATTTTTGTAATTAGGAACAGA 838

Qy     301 CGAGCCATACTGAATAATTTTGTAATCTAACAACAGATGCAATAGTTTAAATAAACTTAT 360
             |     | ||| || || | |   |  | |||||||    | |  |  || ||    |
Db     839 TTAAATCGAATGATTAGTTGTTTTTTTAATCAACAGAACTTACAAATCGAAAAAGTAAAA 898

Qy     361 ACTTAACTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 406
           |   ||     |||||||||||||||||||||||||||||||||||||
Db     899 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 944


RESULT 9
US-08-908-909-5
; Sequence 5, Application US/08908909
; Patent No. 6416960
;  GENERAL INFORMATION:
;    APPLICANT:  Bryan, Bruce
;    TITLE OF INVENTION:  DETECTION AND VISUALIZATION OF
;    TITLE OF INVENTION:  NEOPLASTIC TISSUES AND OTHER TISSUES
;    NUMBER OF SEQUENCES:  14
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brown, Martin, Haller & McClain
;      STREET:  1660 Union Street
;      CITY:  San Diego
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92101-2926
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/908,909
;      FILING DATE:  08-AUG-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/023,374
;      FILING DATE:  08-AUG-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seidman, Stephanie L
;      REGISTRATION NUMBER:  33,779
;      REFERENCE/DOCKET NUMBER:  6680-108
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-238-0999
;      TELEFAX:  619-238-0062
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  958 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FRAGMENT TYPE:
;    ORIGINAL SOURCE:
;    FEATURE:
;      NAME/KEY:  Coding Sequence
;      LOCATION:  115...702
;      OTHER INFORMATION:  apoaequorin-encoding gene
;    PUBLICATION INFORMATION:
;      AUTHORS:  Inouye et al.
;      JOURNAL:  Proc. Natl. Acad. Sci. U.S.A.
;      VOLUME:  82
;      PAGES:  3154-3158
;      DATE:  (1985)
;      DOCUMENT NUMBER:  5,093,240
US-08-908-909-5

  Query Match             12.9%;  Score 52.4;  DB 4;  Length 958;
  Best Local Similarity   57.2%;  Pred. No. 0.0072;
  Matches   95;  Conservative    0;  Mismatches   71;  Indels    0;  Gaps    0;

Qy     241 ATTTGCCATAATGGAGACCCGTTTTTATTGAATATCGTCAGTTTCCAATTAAAGTCATTT 300
           |||| ||| | |   || |  |||  || |  | |  | | |||   | | | |
Db     779 ATTTTCCAAATTTTTGAACGATTTCAATCGTTTGTGTTGATTTTTGTAATTAGGAACAGA 838

Qy     301 CGAGCCATACTGAATAATTTTGTAATCTAACAACAGATGCAATAGTTTAAATAAACTTAT 360
             |     | ||| || || | |   |  | |||||||    | |  |  || ||    |
Db     839 TTAAATCGAATGATTAGTTGTTTTTTTAATCAACAGAACTTACAAATCGAAAAAGTAAAA 898

Qy     361 ACTTAACTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 406
           |   ||     |||||||||||||||||||||||||||||||||||||
Db     899 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 944


RESULT 10
US-09-609-161B-5
; Sequence 5, Application US/09609161B
; Patent No. 6436682
; GENERAL INFORMATION:
;  APPLICANT: Bryan, Bruce
;  APPLICANT:  Szent-Gyorgyi, Christopher
;  APPLICANT:  PROLUME, LTD.
;  TITLE OF INVENTION: LUCIFERASES, FLUORESCENT PROTEINS, NUCLEIC ACIDS ENCODING THE
;  TITLE OF INVENTION:  AND FLUORESCENT PROTEINS AND THE USE THEREOF IN DIAGNOSTICS,
;  TITLE OF INVENTION:  SCREENING AND NOVELTY ITEMS
;  FILE REFERENCE: 24729-121B
;  CURRENT APPLICATION NUMBER: US/09/609,161B
;  CURRENT FILING DATE:  2000-06-30
;  PRIOR APPLICATION NUMBER: 09/277,716
;  PRIOR FILING DATE: 1999-03-26
;  PRIOR APPLICATION NUMBER: 60/102,939
;  PRIOR FILING DATE: 1998-10-01
;  PRIOR APPLICATION NUMBER: 60/089,367
;  PRIOR FILING DATE: 1998-06-15
;  PRIOR APPLICATION NUMBER: 60/079,624
;  PRIOR FILING DATE: 1998-03-27
;  NUMBER OF SEQ ID NOS: 32
;  SOFTWARE: PatentIn Ver. 2.0
```

; SEQ ID NO 5
; LENGTH: 958
; TYPE: DNA
; ORGANISM: Aequorea (luminescent jellyfish)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (115)..(702)
; OTHER INFORMATION: Apoaequorin-encoding gene
; PUBLICATION INFORMATION:
; PATENT DOCUMENT NUMBER: US 07/105,602
; PATENT FILING DATE: 1987-10-08
; PUBLICATION DATE: 1992-03-03
; PUBLICATION INFORMATION:
; AUTHORS: Inouye, S.
; AUTHORS: No. 6436682uchi, M.
; AUTHORS: Sakaki, Y.
; AUTHORS: Takagi, Y.
; AUTHORS: Miyata, T.
; AUTHORS: Iwanaga, S.
; AUTHORS: Miyata, T.
; AUTHORS: Tsuji, F.I.
; TITLE: Cloning and sequence analysis of cDNA for the luminescent protein
; TITLE: aequorin
; JOURNAL: Proc. Natl. Acad. Sci. USA
; VOLUME: 82(10)
; PAGES: 3154-3158
; DATE: 1985-05
US-09-609-161B-5

Query Match 12.9%; Score 52.4; DB 4; Length 958;
Best Local Similarity 57.2%; Pred. No. 0.0072;
Matches 95; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

```
Qy    241 ATTTGCCATAATGGAGACCCGTTTTTATTGAATATCGTCAGTTTCCAATTAAAGTCATTT 300
          |||| ||| | |   || |  |||  || |  | |  | | |||   | | | |
Db    779 ATTTTCCAAATTTTTGAACGATTTCAATCGTTTGTGTTGATTTTTGTAATTAGGAACAGA 838

Qy    301 CGAGCCATACTGAATAATTTTGTAATCTAACAACAGATGCAATAGTTTAAATAAACTTAT 360
            |     | ||| || || | |   |  | |||||||    | |  |  || ||    |
Db    839 TTAAATCGAATGATTAGTTGTTTTTTTAATCAACAGAACTTACAAATCGAAAAAGTAAAA 898

Qy    361 ACTTAACTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 406
          |   ||     |||||||||||||||||||||||||||||||||||
Db    899 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 944
```

RESULT 11
US-08-990-103-5
; Sequence 5, Application US/08990103
; Patent No. 6458547
; GENERAL INFORMATION:
; APPLICANT: Bruce J. Bryan
; APPLICANT: Stephen Gaalema
; APPLICANT: Randall B. Murphy
; TITLE OF INVENTION: APPARATUS AND METHOD FOR DETECTING AND
; TITLE OF INVENTION: IDENTIFYING INFECTIOUS AGENTS
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92101-2926
; COMPUTER READABLE FORM:
; MEDIUM TYPE:
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/990,103
; FILING DATE: 12-DEC-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/037,675,
; FILING DATE: 02-FEB-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/033,745
; FILING DATE: 12-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 6680-112
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-238-0999
; TELEFAX: 619-238-0062
; TELEX:
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 958 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 115...702
; OTHER INFORMATION: apoaequorin-encoding gene
; PUBLICATION INFORMATION:
; AUTHORS: Inouye et al.
; JOURNAL: Proc. Natl. Acad. Sci. U.S.A.
; VOLUME: 82
; PAGES: 3154-3158
; DATE: (1985)
; DOCUMENT NUMBER: 5,093,240
US-08-990-103-5

Query Match 12.9%; Score 52.4; DB 4; Length 958;
Best Local Similarity 57.2%; Pred. No. 0.0072;
Matches 95; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

```
Qy    241 ATTTGCCATAATGGAGACCCGTTTTTATTGAATATCGTCAGTTTCCAATTAAAGTCATTT 300
          |||| ||| | |   || |  |||  || |  | |  | | |||   | | | |
Db    779 ATTTTCCAAATTTTTGAACGATTTCAATCGTTTGTGTTGATTTTTGTAATTAGGAACAGA 838

Qy    301 CGAGCCATACTGAATAATTTTGTAATCTAACAACAGATGCAATAGTTTAAATAAACTTAT 360
            |     | ||| || || | |   |  | |||||||    | |  |  || ||    |
Db    839 TTAAATCGAATGATTAGTTGTTTTTTTAATCAACAGAACTTACAAATCGAAAAAGTAAAA 898

Qy    361 ACTTAACTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 406
          |   ||     |||||||||||||||||||||||||||||||||||
Db    899 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 944
```

RESULT 12
US-08-308-883-1
; Sequence 1, Application US/08308883
; Patent No. 5576300
; GENERAL INFORMATION:
; APPLICANT: Mukerji, P.
; APPLICANT: Prieto, P. A.
; APPLICANT: Seo, A. E.-Y.
; APPLICANT: Baxter, J. H.
; APPLICANT: Cummings, R.D.
; TITLE OF INVENTION: Method for Inhibition of Human Rotavirus Infection.
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lonnie R. Drayer
; ADDRESSEE: ROSS Products Division
; ADDRESSEE: Abbott Laboratories
; STREET: 625 Cleveland Avenue
; CITY: Columbus

```
;       STATE:  Ohio
;       COUNTRY:  United States
;       ZIP:  43215
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  3.5 inch, 1.44 Mb storage
;       COMPUTER:  Apple Macintosh
;       OPERATING SYSTEM:  MacIntosh System 7.1
;       SOFTWARE:  ClarisWorks 1.0
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/308,883
;       FILING DATE:  16-SEP-1994
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:  No. 5576300 applicable
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (614) 624-3774
;       TELEFAX:  (614) 624-3074
;       TELEX:  No. 5576300e
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  857 base pairs
;       TYPE:  Nucleic acid
;       STRANDEDNESS:  Single
;       TOPOLOGY:  Linear
;     MOLECULE TYPE:  cDNA
;       DESCRIPTION:  Human milk kappa-casein
;     HYPOTHETICAL:  No
;     ANTI-SENSE:
;     FRAGMENT TYPE:
;     ORIGINAL SOURCE:  Human
;       ORGANISM:  Homo sapiens
;       STRAIN:
;       INDIVIDUAL ISOLATE:
;       DEVELOPMENTAL STAGE:  Adult
;       HAPLOTYPE:
;       TISSUE TYPE:  Mammary gland
;       CELL TYPE:
;       CELL LINE:
;       ORGANELLE:
;     IMMEDIATE SOURCE:  Human Mammary Gland
;       LIBRARY:
;       CLONE:
;     POSITION IN GENOME:
;       CHROMOSOME/SEGMENT:
;       MAP POSITION:
;       UNITS:
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  45...593
;       IDENTIFICATION METHOD:  DNA sequencing and restriction analysis
;       OTHER INFORMATION:  The encoded product of nucleotide SEQ ID NO: 1:  is the human
;     PUBLICATION INFORMATION:
;       AUTHORS:  L.  Hansson et al
;       TITLE:  DNA Encoding Kappa-Casein,  Process for Obtaining the Protein and Use The
;       JOURNAL:
;       VOLUME:
;       ISSUE:
;       PAGES:
;       DATE:
;       DOCUMENT NUMBER:  PCT/WO93/15196
;       FILING DATE:  25-JAN-1993
;       PUBLICATION DATE:  05-AUG-1993
;       RELEVANT RESIDUES IN SEQ ID NO:
US-08-308-883-1

  Query Match             12.8%;  Score 52;  DB 1;  Length 857;
  Best Local Similarity   67.6%;  Pred. No. 0.0086;
  Matches   73;  Conservative    0;  Mismatches   35;  Indels    0;  Gaps    0;

Qy     299 TTCGAGCCATACTGAATAATTTTGTAATCTAACAACAGATGCAATAGTTTAAATAAACTT 358
           |||  |||| | | |||| |||  |  |   |||| | |  |||  | ||  | |||
Db     741 TTCATGCCACATTCAATATTTTGATTCTTGCACAATAAAGCCAACTGATTGCAAAAAAAA 800

Qy     359 ATACTTAACTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 406
           | |   ||     |||||||||||||||||||||||||||||||||
Db     801 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 848


RESULT 13
US-08-730-163-1
; Sequence 1, Application US/08730163
; Patent No. 5712250
;  GENERAL INFORMATION:
;    APPLICANT:  Mukerji, P.
;    APPLICANT:  Prieto, P. A.
;    APPLICANT:  Seo, A. E.-Y.
;    APPLICANT:  Baxter,  J. H.
;    APPLICANT:  Cummings,  R.D.
;    TITLE OF INVENTION:  Product for Inhibition of Human Rotavirus Infection.
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lonnie R.  Drayer
;      ADDRESSEE:  ROSS Products Division
;      ADDRESSEE:  Abbott Laboratories
;      STREET:  625 Cleveland Avenue
;      CITY:  Columbus
;      STATE:  Ohio
;      COUNTRY:  United States
;      ZIP:  43215
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch, 1.44 Mb storage (B)COMPUTER: Apple Macintosh
;      OPERATING SYSTEM:  MacIntosh System 7.1(D)SOFTWARE: ClarisWorks 1.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/730,163
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/308,882
;      FILING DATE:  16-SEP-1994
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (614) 624-3774
;      TELEFAX:  (614) 624-3074
;      TELEX:  No. 5712250e
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  857 base pairs
;      TYPE:  Nucleic acid
;      STRANDEDNESS:  Single
;      TOPOLOGY:  Linear
;    MOLECULE TYPE:  cDNA
;      DESCRIPTION:  Human milk kappa-casein
;    HYPOTHETICAL:  No
;    ANTI-SENSE:
;    FRAGMENT TYPE:
;    ORIGINAL SOURCE:  Human
;      ORGANISM:  Homo sapiens
;      STRAIN:
;      INDIVIDUAL ISOLATE:
;      DEVELOPMENTAL STAGE:  Adult
;      HAPLOTYPE:
;      TISSUE TYPE:  Mammary gland
;      CELL TYPE:
;      CELL LINE:
;      ORGANELLE:
;    IMMEDIATE SOURCE:  Human Mammary Gland
;      LIBRARY:
;      CLONE:
;    POSITION IN GENOME:
;      CHROMOSOME/SEGMENT:
;      MAP POSITION:
;      UNITS:
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  45...593
;      IDENTIFICATION METHOD:  DNA sequencing and restriction analysis
;      OTHER INFORMATION:  The encoded product of nucleotide SEQ ID NO: 1:  is the hu
;    PUBLICATION INFORMATION:
```

```
;      AUTHORS:  L.  Hansson et al
;      TITLE:  DNA Encoding Kappa-Casein,  Process for Obtaining the Protein and Use The
;      JOURNAL:
;      VOLUME:
;      ISSUE:
;      PAGES:
;      DATE:
;      DOCUMENT NUMBER:  PCT/WO93/15196
;      FILING DATE:  25-JAN-1993
;      PUBLICATION DATE:  05-AUG-1993
;      RELEVANT RESIDUES IN SEQ ID NO:
US-08-730-163-1

  Query Match             12.8%;  Score 52;  DB 1;  Length 857;
  Best Local Similarity   67.6%;  Pred. No. 0.0086;
  Matches   73;  Conservative    0;  Mismatches   35;  Indels    0;  Gaps    0;

Qy     299 TTCGAGCCATACTGAATAATTTTGTAATCTAACAACAGATGCAATAGTTTAAATAAACTT 358
           |||  |||| | | |||| |||  |  |   |||| | |  |||  | ||  | |||
Db     741 TTCATGCCACATTCAATATTTTGATTCTTGCACAATAAAGCCAACTGATTGCAAAAAAAA 800

Qy     359 ATACTTAACTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 406
           | |   ||     ||||||||||||||||||||||||||||||||||||||||||
Db     801 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 848


RESULT 14
US-08-256-799-1
; Sequence 1, Application US/08256799
; Patent No. 6222094
;  GENERAL INFORMATION:
;    APPLICANT:  HANSSON, Lennart
;    APPLICANT:  STROEMQVIST, Mats
;    APPLICANT:  BERGSTROEM, Sven
;    APPLICANT:  HERNELL, Olle
;    APPLICANT:  Toernell, Jan
;    TITLE OF INVENTION:  DNA ENCODING KAPPA-CASEIN, PROCESS FOR
;    TITLE OF INVENTION:  OBTAINING THE PROTEIN AND USE THEREOF
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/256,799
;      FILING DATE:  06-DEC-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DK 88/92
;      FILING DATE:  23-JAN-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  COOPER, Iver P.
;      REGISTRATION NUMBER:  28,005
;      REFERENCE/DOCKET NUMBER:  HANSSON=1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  857 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  45..593
;    FEATURE:
;      NAME/KEY:  mat_peptide
;      LOCATION:  45..593
;    FEATURE:
;      NAME/KEY:  sig_peptide
;      LOCATION:  45..104
;    FEATURE:
;      NAME/KEY:  5'UTR
;      LOCATION:  13..44
;    FEATURE:
;      NAME/KEY:  3'UTR
;      LOCATION:  594..848
US-08-256-799-1

  Query Match             12.8%;  Score 52;  DB 4;  Length 857;
  Best Local Similarity   67.6%;  Pred. No. 0.0086;
  Matches   73;  Conservative    0;  Mismatches   35;  Indels    0;  Gaps    0;

Qy     299 TTCGAGCCATACTGAATAATTTTGTAATCTAACAACAGATGCAATAGTTTAAATAAACTT 358
           |||  |||| | | |||| |||  |  |   |||| | |  |||  | ||  | |||
Db     741 TTCATGCCACATTCAATATTTTGATTCTTGCACAATAAAGCCAACTGATTGCAAAAAAAA 800

Qy     359 ATACTTAACTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 406
           | |   ||     ||||||||||||||||||||||||||||||||||||||||||
Db     801 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 848


RESULT 15
US-08-462-437-1
; Sequence 1, Application US/08462437
; Patent No. 6232094
;  GENERAL INFORMATION:
;    APPLICANT:  HANSSON, Lennart
;    APPLICANT:  STROEMQVIST, Mats
;    APPLICANT:  BERGSTROEM, Sven
;    APPLICANT:  HERNELL, Olle
;    APPLICANT:  TOERNELL, Jan
;    TITLE OF INVENTION:  DNA ENCODING KAPPA-CASEIN, PROCESS
;    TITLE OF INVENTION:  FOR OBTAINING THE PROTEIN AND USE THEREOF
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BROWDY AND NEIMARK
;      STREET:  419 Seventh Street, N.W., Suite 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/462,437
;      FILING DATE:  05-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DK 88/92
;      FILING DATE:  23-JAN-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  COOPER, Iver P.
;      REGISTRATION NUMBER:  28,005
;      REFERENCE/DOCKET NUMBER:  HANSSON=1A
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-628-5197
;      TELEFAX:  202-737-3528
;      TELEX:  248633
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
```

```
;       LENGTH:  857 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     HYPOTHETICAL:  NO
;     ORIGINAL SOURCE:
;       ORGANISM:  Homo sapiens
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  45..593
;     FEATURE:
;       NAME/KEY:  mat_peptide
;       LOCATION:  45..593
;     FEATURE:
;       NAME/KEY:  sig_peptide
;       LOCATION:  45..104
;     FEATURE:
;       NAME/KEY:  5'UTR
;       LOCATION:  13..44
;     FEATURE:
;       NAME/KEY:  3'UTR
;       LOCATION:  594..848
US-08-462-437-1

  Query Match             12.8%;  Score 52;  DB 4;  Length 857;
  Best Local Similarity   67.6%;  Pred. No. 0.0086;
  Matches   73;  Conservative    0;  Mismatches   35;  Indels    0;  Gaps    0;

Qy     299 TTCGAGCCATACTGAATAATTTTGTAATCTAACAACAGATGCAATAGTTTAAATAAACTT 358
           |||  |||| | | |||| |||  |  |   |||| | |  |||  | ||  | |||
Db     741 TTCATGCCACATTCAATATTTTGATTCTTGCACAATAAAGCCAACTGATTGCAAAAAAAA 800

Qy     359 ATACTTAACTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 406
           | |   ||     ||||||||||||||||||||||||||||||||||||
Db     801 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 848


Search completed: May  5, 2003, 20:26:55
Job time : 54 secs
```

OM nucleic - nucleic search, using sw model

Run on: May 5, 2003, 19:53:13 ; Search time 143 Seconds
(without alignments)
3354.812 Million cell updates/sec

Title: US-09-829-481-3
Perfect score: 406
Sequence: 1 ctctactacaatcactaagt.....aaaaaaaaaaaaaaaaaaa 406

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 746064 seqs, 590810554 residues

Total number of hits satisfying chosen parameters: 1492128

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	406	100.0	406	10	US-09-829-481-3	Sequence 3, Appli
	2	76	18.7	386	10	US-09-829-481-5	Sequence 5, Appli
	3	55.2	13.6	325	9	US-10-091-483-30	Sequence 30, Appl
	4	55.2	13.6	325	10	US-09-764-846-30	Sequence 30, Appl
	5	55.2	13.6	621	9	US-10-091-483-105	Sequence 105, App
	6	55.2	13.6	621	10	US-09-764-846-105	Sequence 105, App
	7	54.6	13.4	351	10	US-09-829-481-7	Sequence 7, Appli
c	8	54.6	13.4	425	10	US-09-834-975-451	Sequence 451, App
c	9	53.8	13.3	380	10	US-09-960-352-9335	Sequence 9335, Ap
	10	53.8	13.3	461	10	US-09-829-481-1	Sequence 1, Appli
	11	53.6	13.2	471	9	US-09-918-995-14052	Sequence 14052, A
c	12	53.2	13.1	236	10	US-09-960-352-12183	Sequence 12183, A
c	13	53	13.1	397	10	US-09-960-352-13784	Sequence 13784, A
	14	52.8	13.0	393	10	US-09-960-352-5187	Sequence 5187, Ap
	15	52.8	13.0	454	9	US-09-918-995-13971	Sequence 13971, A
	16	52.6	13.0	3293	10	US-09-764-864-123	Sequence 123, App
	17	52.4	12.9	958	9	US-10-126-139-5	Sequence 5, Appli
	18	52.4	12.9	958	9	US-10-126-798-5	Sequence 5, Appli
	19	52.4	12.9	958	10	US-09-803-211-5	Sequence 5, Appli
	20	52.4	12.9	958	10	US-09-746-485A-5	Sequence 5, Appli
c	21	52.2	12.9	368	10	US-09-834-975-47	Sequence 47, Appl
	22	52.2	12.9	426	9	US-09-918-995-17732	Sequence 17732, A
	23	52.2	12.9	432	9	US-09-918-995-13651	Sequence 13651, A
	24	52.2	12.9	1387	9	US-10-091-483-113	Sequence 113, App
	25	52.2	12.9	1387	10	US-09-764-846-113	Sequence 113, App
	26	52.2	12.9	3957	10	US-09-764-853-77	Sequence 77, Appl
	27	52	12.8	857	9	US-09-954-531-977	Sequence 977, App
	28	52	12.8	1499	9	US-10-114-893-70	Sequence 70, Appl
	29	51.8	12.8	219	10	US-09-878-574-15258	Sequence 15258, A
	30	51.8	12.8	1457	9	US-09-989-920-93	Sequence 93, Appl
	31	51.8	12.8	4000	9	US-09-981-353-8	Sequence 8, Appli
c	32	51.6	12.7	337	10	US-09-960-352-6976	Sequence 6976, Ap
c	33	51.6	12.7	347	10	US-09-960-352-3371	Sequence 3371, Ap
	34	51.4	12.7	1772	9	US-10-102-806-196	Sequence 196, App
c	35	51	12.6	358	10	US-09-960-352-5559	Sequence 5559, Ap
c	36	50.8	12.5	375	10	US-09-960-352-15014	Sequence 15014, A
c	37	50.8	12.5	396	9	US-09-970-966-53	Sequence 53, Appl
c	38	50.8	12.5	396	10	US-09-825-294-53	Sequence 53, Appl
	39	50.8	12.5	1001	9	US-09-986-480-93	Sequence 93, Appl
	40	50.6	12.5	1411	9	US-10-000-273-5	Sequence 5, Appli
	41	50.6	12.5	3239	9	US-10-174-363-39	Sequence 39, Appl
	42	50.4	12.4	2511	9	US-10-091-752A-13	Sequence 13, Appl
	43	50.2	12.4	285	10	US-09-834-975-668	Sequence 668, App
c	44	50.2	12.4	291	10	US-09-960-352-1243	Sequence 1243, Ap
	45	50.2	12.4	373	10	US-09-960-352-7994	Sequence 7994, Ap

ALIGNMENTS

RESULT 1
US-09-829-481-3
; Sequence 3, Application US/09829481
; Patent No. US20020069427A1
; GENERAL INFORMATION:
; APPLICANT: Presnail, James
; APPLICANT: Weng, Zude
; APPLICANT: Wong, James
; TITLE OF INVENTION: Arthropod Defensins
; FILE REFERENCE: BB1441 US NA
; CURRENT APPLICATION NUMBER: US/09/829,481
; CURRENT FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/197279
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 3
; LENGTH: 406
; TYPE: DNA
; ORGANISM: Vaejovis carolinianus
US-09-829-481-3

Query Match 100.0%; Score 406; DB 10; Length 406;
Best Local Similarity 100.0%; Pred. No. 1.8e-74;
Matches 406; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 CTCTACTACAATCACTAAGTTCTTTCTCCACTCAGCTTCAAGAATGAAATCCATAGCTAT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CTCTACTACAATCACTAAGTTCTTTCTCCACTCAGCTTCAAGAATGAAATCCATAGCTAT 60

Qy      61 TATTTTCATCGTTCTTGTTGCCTTCTGTATTTTGGAGGATGGGATTGTAGAAGCTGGTTT 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TATTTTCATCGTTCTTGTTGCCTTCTGTATTTTGGAGGATGGGATTGTAGAAGCTGGTTT 120

Qy     121 TGGATGTCCCTTTAATGCAGGAAAATGCCATAGACATTGCAAAAGTATTCGTCGTAGAGG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 TGGATGTCCCTTTAATGCAGGAAAATGCCATAGACATTGCAAAAGTATTCGTCGTAGAGG 180

Qy     181 AGGCTTTTGCAGAGGAACTTTCAGGACAACCTGCGTTTGCTATAGGTGAAAATCCGATTT 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 AGGCTTTTGCAGAGGAACTTTCAGGACAACCTGCGTTTGCTATAGGTGAAAATCCGATTT 240
```

```
Qy     241 ATTTGCCATAATGGAGACCCGTTTTTATTGAATATCGTCAGTTTCCAATTAAAGTCATTT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 ATTTGCCATAATGGAGACCCGTTTTTATTGAATATCGTCAGTTTCCAATTAAAGTCATTT 300

Qy     301 CGAGCCATACTGAATAATTTTGTAATCTAACAACAGATGCAATAGTTTAAATAAACTTAT 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CGAGCCATACTGAATAATTTTGTAATCTAACAACAGATGCAATAGTTTAAATAAACTTAT 360

Qy     361 ACTTAACTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 406
           ||||||||||||||||||||||||||||||||||||||||||||||
Db     361 ACTTAACTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 406


RESULT 2
US-09-829-481-5
; Sequence 5, Application US/09829481
; Patent No. US20020069427A1
; GENERAL INFORMATION:
;  APPLICANT: Presnail, James
;  APPLICANT:  Weng, Zude
;  APPLICANT:  Wong, James
;  TITLE OF INVENTION: Arthropod Defensins
;  FILE REFERENCE: BB1441 US NA
;  CURRENT APPLICATION NUMBER: US/09/829,481
;  CURRENT FILING DATE:  2001-04-10
;  PRIOR APPLICATION NUMBER: 60/197279
;  PRIOR FILING DATE: 2000-04-14
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
;   LENGTH: 386
;   TYPE: DNA
;   ORGANISM: Argiope sp.
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (351)
;   NAME/KEY: unsure
;   LOCATION: (364)..(365)
;   NAME/KEY: unsure
;   LOCATION: (386)
US-09-829-481-5

  Query Match             18.7%;  Score 76;  DB 10;  Length 386;
  Best Local Similarity   54.4%;  Pred. No. 9e-07;
  Matches  178;  Conservative    0;  Mismatches  141;  Indels    8;  Gaps    1;

Qy      78 TTGCCTTCTGTATTTTGGAGGATGGGATTGTAGAAGCTGGTTTTGGATGTCCCTTTAATG 137
           | | | |||||  ||| |     | |   || ||||||||||| || || |||||  |
Db      45 TAGTCGTCTGTGCTTTTGCCACAGTGGCCGTGGAAGCTGGTTTCGGCTGCCCCTTCGACC 104

Qy     138 CAGGAAAATGCCATAGACATTGCAAAAGTATTCGTCGTAGAGGAGGCTTTTGCAGAGGAA 197
             | || || |  |||||||  || ||       || ||||| |  ||||
Db     105 AGATGCAGTGTCACAATCATTGCAGGAGCATCAAATACAGGGGAGGATACTGCACCAACT 164

Qy     198 CTTTCAGGACAACCTGCGTTTGCTATAGGTGAAAATCCGATTTATTTGCCATAATGGAGA 257
           ||||  |   ||||||   || ||  | |||  | ||   |    |  || |  ||||
Db     165 TATTCAAGCGCACCTGCAAGTGTTACGGATGATGACCCCCCTCCCCTCTCAGACAGGAGC 224

Qy     258 CCCGTTTTTATTGAATATCGTCAGTTTCCAATTAAAGTCATTTCGAGCCATACTGAATAA 317
           ||     ||        || |     ||| |||    |  |    |  || | || |  |
Db     225 CCACACCTT--------TCATTGACATCCGATTCCGATTTTCCAAATGCAAATTGTAACA 276

Qy     318 TTTTGTAATCTAACAACAGATGCAATAGTTTAAATAAACTTATACTTAACTTTTAAAAAA 377
             |  | |  |     ||  |||  ||  || ||  || ||  | || || ||| |||||
Db     277 CATGATGAATTTGATGCAAGTGCCTTAAATTTAAATAAATTTGATTTTACATTTTAAAAA 336

Qy     378 AAAAAAAAAAAAAAAAAAAAAAAAAA 404
           |||||||||||||| ||||||||||||
Db     337 AAAAAAAAAAAAAAANAAAAAAAAAAA 363


RESULT 3
US-10-091-483-30
; Sequence 30, Application US/10091483
; Publication No. US20030049650A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PTZ12C1
;  CURRENT APPLICATION NUMBER: US/10/091,483
;  CURRENT FILING DATE:  2002-03-07
;  NUMBER OF SEQ ID NOS: 348
;  Prior Application removed - See File Wrapper or Palm
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
;   LENGTH: 325
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-091-483-30

  Query Match             13.6%;  Score 55.2;  DB 9;  Length 325;
  Best Local Similarity   60.8%;  Pred. No. 0.016;
  Matches   90;  Conservative    0;  Mismatches   58;  Indels    0;  Gaps    0;

Qy     259 CCGTTTTTATTGAATATCGTCAGTTTCCAATTAAAGTCATTTCGAGCCATACTGAATAAT 318
           || || | ||| | ||   | |     | ||||  |||   |  ||  ||    ||| ||
Db     104 CCTTTATAATTCACTAAAATAAAGCATCTATTAGTGTCTGATTTAGGAATGTAAAATGAT 163

Qy     319 TTTGTAATCTAACAACAGATGCAATAGTTTAAATAAACTTATACTTAACTTTTAAAAAAA 378
           | |||| |     ||   |    ||   ||  | |||  |||    ||| | ||||||| 
Db     164 TCTGTATTAATGTAAATAAGATTATCTATTGCAAAAAGATATTTCAAACCTAAAAAAAAA 223

Qy     379 AAAAAAAAAAAAAAAAAAAAAAAAAAA 406
           |||||||||||||||||||||||||||
Db     224 AAAAAAAAAAAAAAAAAAAAAAAAAAA 251


RESULT 4
US-09-764-846-30
; Sequence 30, Application US/09764846
; Patent No. US20020102638A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PTZ12
;  CURRENT APPLICATION NUMBER: US/09/764,846
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 348
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
;   LENGTH: 325
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-846-30

  Query Match             13.6%;  Score 55.2;  DB 10;  Length 325;
  Best Local Similarity   60.8%;  Pred. No. 0.016;
  Matches   90;  Conservative    0;  Mismatches   58;  Indels    0;  Gaps    0;

Qy     259 CCGTTTTTATTGAATATCGTCAGTTTCCAATTAAAGTCATTTCGAGCCATACTGAATAAT 318
           || || | ||| | ||   | |     | ||||  |||   |  ||  ||    ||| ||
Db     104 CCTTTATAATTCACTAAAATAAAGCATCTATTAGTGTCTGATTTAGGAATGTAAAATGAT 163

Qy     319 TTTGTAATCTAACAACAGATGCAATAGTTTAAATAAACTTATACTTAACTTTTAAAAAAA 378
           | |||| |     ||   |    ||   ||  | |||  |||    ||| | ||||||| 
Db     164 TCTGTATTAATGTAAATAAGATTATCTATTGCAAAAAGATATTTCAAACCTAAAAAAAAA 223

Qy     379 AAAAAAAAAAAAAAAAAAAAAAAAAAA 406
           |||||||||||||||||||||||||||
Db     224 AAAAAAAAAAAAAAAAAAAAAAAAAAA 251


RESULT 5
US-10-091-483-105
; Sequence 105, Application US/10091483
```

```
; Publication No. US20030049650A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PTZ12C1
;  CURRENT APPLICATION NUMBER: US/10/091,483
;  CURRENT FILING DATE:  2002-03-07
;  NUMBER OF SEQ ID NOS: 348
;  Prior Application removed - See File Wrapper or Palm
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 105
;   LENGTH: 621
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (80)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (612)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (620)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc_feature
;   LOCATION: (621)
;   OTHER INFORMATION: n equals a,t,g, or c
US-10-091-483-105

  Query Match             13.6%;  Score 55.2;  DB 9;  Length 621;
  Best Local Similarity   60.8%;  Pred. No. 0.02;
  Matches   90;  Conservative    0;  Mismatches   58;  Indels    0;  Gaps    0;

Qy     259 CCGTTTTTATTGAATATCGTCAGTTTCCAATTAAAGTCATTTCGAGCCATACTGAATAAT 318
           || || | ||| | ||   | |     | ||||  |||   |  ||  ||    ||| ||
Db     388 CCTTTATAATTCACTAAAATAAAGCATCTATTAGTGTCTGATTTAGGAATGTAAAATGAT 447

Qy     319 TTTGTAATCTAACAACAGATGCAATAGTTTAAATAAACTTATACTTAACTTTTAAAAAAA 378
           | |||| |     ||   |    ||   ||  | |||  |||    ||| |  |||||||
Db     448 TCTGTATTAATGTAAATAAGATTATCTATTGCAAAAAGATATTTCAAACCTAAAAAAAAA 507

Qy     379 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 406
           ||||||||||||||||||||||||||||||
Db     508 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 535


RESULT 6
US-09-764-846-105
; Sequence 105, Application US/09764846
; Patent No. US20020102638A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PTZ12
;  CURRENT APPLICATION NUMBER: US/09/764,846
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - consult PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 348
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 105
;   LENGTH: 621
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (80)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: SITE
;   LOCATION: (612)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: SITE
;   LOCATION: (620)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: SITE
;   LOCATION: (621)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-764-846-105

  Query Match             13.6%;  Score 55.2;  DB 10;  Length 621;
  Best Local Similarity   60.8%;  Pred. No. 0.02;
  Matches   90;  Conservative    0;  Mismatches   58;  Indels    0;  Gaps    0;

Qy     259 CCGTTTTTATTGAATATCGTCAGTTTCCAATTAAAGTCATTTCGAGCCATACTGAATAAT 318
           || || | ||| | ||   | |     | ||||  |||   |  ||  ||    ||| ||
Db     388 CCTTTATAATTCACTAAAATAAAGCATCTATTAGTGTCTGATTTAGGAATGTAAAATGAT 447

Qy     319 TTTGTAATCTAACAACAGATGCAATAGTTTAAATAAACTTATACTTAACTTTTAAAAAAA 378
           | |||| |     ||   |    ||   ||  | |||  |||    ||| |  |||||||
Db     448 TCTGTATTAATGTAAATAAGATTATCTATTGCAAAAAGATATTTCAAACCTAAAAAAAAA 507

Qy     379 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 406
           ||||||||||||||||||||||||||||||
Db     508 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 535


RESULT 7
US-09-829-481-7
; Sequence 7, Application US/09829481
; Patent No. US20020069427A1
; GENERAL INFORMATION:
;  APPLICANT: Presnail, James
;  APPLICANT:  Weng, Zude
;  APPLICANT:  Wong, James
;  TITLE OF INVENTION: Arthropod Defensins
;  FILE REFERENCE: BB1441 US NA
;  CURRENT APPLICATION NUMBER: US/09/829,481
;  CURRENT FILING DATE:  2001-04-10
;  PRIOR APPLICATION NUMBER: 60/197279
;  PRIOR FILING DATE: 2000-04-14
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
;   LENGTH: 351
;   TYPE: DNA
;   ORGANISM: Argiope sp.
US-09-829-481-7

  Query Match             13.4%;  Score 54.6;  DB 10;  Length 351;
  Best Local Similarity   54.9%;  Pred. No. 0.021;
  Matches  130;  Conservative    0;  Mismatches  104;  Indels    3;  Gaps    1;

Qy      23 TTTCTCCACTCAGCTTCAAGAATGAAATCCATAGCTATTATTTTCATCGTTCTTGTTGCC 82
           | | || ||    |   || ||||||  | | || | |  | || |||   || ||   |
Db       7 TCTGTCGACATTTCCAAAAAAATGAATGCGAGAGTTCTGTTGTTGATCTGCCTAGT---C 63

Qy      83 TTCTGTATTTTGGAGGATGGGATTGTAGAAGCTGGTTTTGGATGTCCCTTTAATGCAGGA 142
           |||||  ||| |     | ||  || ||||||||||| || || |||||  |
Db      64 GTCTGTGCTTTTGCCACAGTGACCGTGGAAGCTGGTTTCGGCTGCCCCTTCGACCAGATG 123

Qy     143 AAATGCCATAGACATTGCAAAAGTATTCGTCGTAGAGGAGGCTTTTGCAGAGGAACTTTC 202
           | || || |  |||||||  || ||      ||| ||||| |  ||||        |||
Db     124 CAGTGTCACAATCATTGCAGGAGCATCAAATATAGGGGAGGATACTGCACCAACTTATTC 183

Qy     203 AGGACAACCTGCGTTTGCTATAGGTGAAAATCCGATTTATTTGCCATAATGGAGACC 259
           | |   ||||||   || ||  | |||  | ||    |    |  || |  |||| ||
Db     184 AAGCGCACCTGCAAGTGTTACGGATGATGACCCCCCTCCCCTCTCAGACAGGAGCCC 240


RESULT 8
US-09-834-975-451/c
; Sequence 451, Application US/09834975
; Patent No. US20020110815A1
; GENERAL INFORMATION:
;  APPLICANT: Lillie, James
;  APPLICANT:  Brown, Jeffrey
;  APPLICANT:  Bolt, Andrew
```

```
;  APPLICANT:  Van Huffel, Christophe
;  TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS AND METHODS
;  TITLE OF INVENTION:  FOR THE IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
;  TITLE OF INVENTION:  OF HUMAN CANCERS
;  FILE REFERENCE: MRI-016B
;  CURRENT APPLICATION NUMBER: US/09/834,975
;  CURRENT FILING DATE:  2001-04-13
;  PRIOR APPLICATION NUMBER: 60/197,538
;  PRIOR FILING DATE: 2000-04-14
;  NUMBER OF SEQ ID NOS: 1046
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 451
;   LENGTH: 425
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(425)
;   OTHER INFORMATION: n = A,T,C or G
US-09-834-975-451

  Query Match             13.4%;  Score 54.6;  DB 10;  Length 425;
  Best Local Similarity   53.4%;  Pred. No. 0.023;
  Matches  111;  Conservative    0;  Mismatches   97;  Indels    0;  Gaps    0;

Qy    199 TTTCAGGACAACCTGCGTTTGCTATAGGTGAAAATCCGATTTATTTGCCATAATGGAGAC 258
          ||| |  | || | | |   |  | |    |||| ||  ||| |||    |   || ||
Db    215 TTTTAAAAAAAACGGGGGGGGGAAAAAAAAAAAAACCCTTTTTTTTTTTTTTTTGGGGGAA 156

Qy    259 CCGTTTTTATTGAATATCGTCAGTTTCCAATTAAAGTCATTTCGAGCCATACTGAATAAT 318
          |||||  ||   | |  |   ||   | ||    |  |||      | |   || ||
Db    155 ATTTTTTTTTTTTTTTTTTTTTTTAAAATTTTTTTTTTTTTTTNNNAAAAAAAAAAAAA 96

Qy    319 TTTGTAATCTAACAACAGATGCAATAGTTTAAATAAACTTATACTTAACTTTTAAAAAAA 378
             ||   || || | |   || |    ||| |||   | |   ||     |||||||
Db     95 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 36

Qy    379 AAAAAAAAAAAAAAAAAAAAAAAAA 406
          |||||||||||||||||||||||||
Db     35 AAAAAAAAAAAAAAAAAAAAAAAAA 8


RESULT 9
US-09-960-352-9335/c
; Sequence 9335, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
;  APPLICANT: Warren, Wesley C.
;  APPLICANT:  Tao, Nengbing
;  APPLICANT:  Byatt, John C.
;  APPLICANT:  Mathialagan, Nagappan
;  TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
;  TITLE OF INVENTION:  MUSCLE AND FAT DEPOSITION
;  FILE REFERENCE: 16511.006/37-21(10298)C
;  CURRENT APPLICATION NUMBER: US/09/960,352
;  CURRENT FILING DATE:  2001-09-24
;  NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 9335
;   LENGTH: 380
;   TYPE: DNA
;   ORGANISM: Bos taurus
;   OTHER INFORMATION: Clone ID: 40-LIB3058-035-Q1-K1-B8
US-09-960-352-9335

  Query Match             13.3%;  Score 53.8;  DB 10;  Length 380;
  Best Local Similarity   60.7%;  Pred. No. 0.032;
  Matches   88;  Conservative    0;  Mismatches   57;  Indels    0;  Gaps    0;

Qy    262 TTTTTATTGAATATCGTCAGTTTCCAATTAAAGTCATTTCGAGCCATACTGAATAATTTT 321
          |||||  || | |||  |   |||    ||||  |  |||  |    |  | || |||| |
Db    207 TTTTTTTTAATTATTCTTTTTTTTTTTTTTAATTTTTTTTAAAAAATTTTTAAACAATTAT 148

Qy    322 GTAATCTAACAACAGATGCAATAGTTTAAATAAACTTATACTTAACTTTTAAAAAAAAAA 381
          |  | |||      |   | || | ||||||||  |     ||  |  | ;|||||||||
Db    147 ATTTTTTAAATTTGAAAATATTATTATAAATAAAAAAAATAAAAAAATAAATAAAAAAAA 88

Qy    382 AAAAAAAAAAAAAAAAAAAAAAAAA 406
          |||||||||||||||||||||||||
Db     87 TAAAAAAAAAAAAAAAAAAAAAAAA 63


RESULT 10
US-09-829-481-1
; Sequence 1, Application US/09829481
; Patent No. US20020069427A1
; GENERAL INFORMATION:
;  APPLICANT: Presnail, James
;  APPLICANT:  Weng, Zude
;  APPLICANT:  Wong, James
;  TITLE OF INVENTION: Arthropod Defensins
;  FILE REFERENCE: BB1441 US NA
;  CURRENT APPLICATION NUMBER: US/09/829,481
;  CURRENT FILING DATE:  2001-04-10
;  PRIOR APPLICATION NUMBER: 60/197279
;  PRIOR FILING DATE: 2000-04-14
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
;   LENGTH: 461
;   TYPE: DNA
;   ORGANISM: Scolopendra canidens DS
;   FEATURE:
;   NAME/KEY: unsure
;   LOCATION: (408)
;   NAME/KEY: unsure
;   LOCATION: (439)
;   NAME/KEY: unsure
;   LOCATION: (455)
US-09-829-481-1

  Query Match             13.3%;  Score 53.8;  DB 10;  Length 461;
  Best Local Similarity   54.2%;  Pred. No. 0.034;
  Matches  109;  Conservative    0;  Mismatches   92;  Indels    0;  Gaps    0;

Qy     44 ATGAAATCCATAGCTATTATTTTCATCGTTCTTGTTGCCTTCTGTATTTTGGAGGATGGG 103
          |||||  |  |   | |||| ||| | |||  | ||    | |     || |  |
Db     50 ATGAAGACTGTGTATGTTATCTTCTTAGTTGCTCTTCTTGTATTAGCATTAGCTGGAACT 109

Qy    104 ATTGTAGAAGCTGGTTTTGGATGTCCCTTTAATGCAGGAAAATGCCATAGACATTGCAAA 163
             || |||||||||| || || |||    |   |    ||||| ||||||||||||
Db    110 TACGTGGAAGCTGGTTTCGGCTGCCCCGGGGACCAATATGAATGCAATAGACATTGCAGG 169

Qy    164 AGTATTCGTCGTAGAGGAGGCTTTTGCAGAGGAACTTTCAGGACAACCTGCGTTTGCTAT 223
          ||| | |   ||  || || |  ||||  |||  | | | |   || ||   |||||
Db    170 GGTAATGGATTTACTGGGGGTTACTGCACTGGATTTCTGAAGTTCACGTGTAAATGCTAC 229

Qy    224 AGGTGAAAATCCGATTTATTT 244
          |  |||  |    ||   |||
Db    230 ACATGATCAAAATATGATTTT 250


RESULT 11
US-09-918-995-14052
; Sequence 14052, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 20411-756
;  CURRENT APPLICATION NUMBER: US/09/918,995
;  CURRENT FILING DATE:  2001-07-30
;  PRIOR APPLICATION NUMBER: US/09/235,076
;  PRIOR FILING DATE: 1999-01-20
;  NUMBER OF SEQ ID NOS: 38054
;  SOFTWARE: FastSEQ for Windows Version 3.0
```

```
; SEQ ID NO 14052
;   LENGTH: 471
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(471)
;   OTHER INFORMATION: n = A,T,C or G
US-09-918-995-14052

  Query Match            13.2%;  Score 53.6;  DB 9;  Length 471;
  Best Local Similarity  59.3%;  Pred. No. 0.038;
  Matches   89;  Conservative    0;  Mismatches   61;  Indels    0;  Gaps    0;

Qy     257 ACCCGTTTTTATTGAATATCGTCAGTTTCCAATTAAAGTCATTTCGAGCCATACTGAATA 316
           || |  ||||   ||| |  |||| |   |   ||| | || |  |   | || | ||||
Db     296 ACACAATTTTGAAGAAAACTGTCAATCAGCTTATAACGACAATGTGGCACTTAATAAATA 355

Qy     317 ATTTTGTAATCTAACAACAGATGCAATAGTTTAAATAAACTTATACTTAACTTTTAAAAA 376
            || |   | ||      | |   || |    ||| |||   | |   ||     |||||
Db     356 CTTGTCAGAACTTTANNAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 415

Qy     377 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 406
           ||||||||||||||||||||||||||||||
Db     416 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 445


RESULT 12
US-09-960-352-12183/c
; Sequence 12183, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
;  APPLICANT: Warren, Wesley C.
;  APPLICANT:  Tao, Nengbing
;  APPLICANT:  Byatt, John C.
;  APPLICANT:  Mathialagan, Nagappan
;  TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
;  TITLE OF INVENTION:  MUSCLE AND FAT DEPOSITION
;  FILE REFERENCE: 16511.006/37-21(10298)C
;  CURRENT APPLICATION NUMBER: US/09/960,352
;  CURRENT FILING DATE:  2001-09-24
;  NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 12183
;   LENGTH: 236
;   TYPE: DNA
;   ORGANISM: Bos taurus
;   OTHER INFORMATION: Clone ID: 52-LIB3058-016-Q1-K1-E12
US-09-960-352-12183

  Query Match            13.1%;  Score 53.2;  DB 10;  Length 236;
  Best Local Similarity  57.1%;  Pred. No. 0.036;
  Matches   97;  Conservative    0;  Mismatches   73;  Indels    0;  Gaps    0;

Qy     237 ATTTATTTGCCATAATGGAGACCCGTTTTTATTGAATATCGTCAGTTTCCAATTAAAGTC 296
           | || ||| ||   ||     || ||   | ||    ||  |||   |   |||    |
Db     220 AATTTTTTTCCGGTATATTTTCCAGTACATTTTTGTCATTTTCACACTAACATTTTTATT 161

Qy     297 ATTTCGAGCCATACTGAATAATTTTTGTAATCTAACAACAGATGCAATAGTTTAAATAAAC 356
            |||       || | | | ||||  || || || || | |   || |    ||| |||
Db     160 TTTTTTTAATTTAATTATTTATTTAATATTCCAAAAAAAAAAAAAAAAAAAAAAAAAAAA 101

Qy     357 TTATACTTAACTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 406
             | |   ||    ||||||||||||||||||||||||||||||||||||||
Db     100 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 51


RESULT 13
US-09-960-352-13784/c
; Sequence 13784, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
;  APPLICANT: Warren, Wesley C.
;  APPLICANT:  Tao, Nengbing
;  APPLICANT:  Byatt, John C.
;  APPLICANT:  Mathialagan, Nagappan
;  TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AN
;  TITLE OF INVENTION:  MUSCLE AND FAT DEPOSITION
;  FILE REFERENCE: 16511.006/37-21(10298)C
;  CURRENT APPLICATION NUMBER: US/09/960,352
;  CURRENT FILING DATE:  2001-09-24
;  NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 13784
;   LENGTH: 397
;   TYPE: DNA
;   ORGANISM: Bos taurus
;   OTHER INFORMATION: Clone ID: 59-LIB3057-006-Q1-K1-G4
US-09-960-352-13784

  Query Match            13.1%;  Score 53;  DB 10;  Length 397;
  Best Local Similarity  55.0%;  Pred. No. 0.048;
  Matches  104;  Conservative    0;  Mismatches   85;  Indels    0;  Gaps    0;

Qy     218 TGCTATAGGTGAAAATCCGATTTATTTGCCATAATGGAGACCCGTTTTTATTGAATATCG 277
           | |  ||  | |||||    ||| ||  |  ||    | |    |||||||  || |
Db     252 TTCAGTATATTAAAATTATTTTTTTTATCATTATAAAAAAAATTTTTTTATAAAACAAAA 193

Qy     278 TCAGTTTCCAATTAAAGTCATTTCGAGCCATACTGAATAATTTTGTAATCTAACAACAGA 337
             | |||   | || |   |||   |   | |   || ||      ||   || || | |
Db     192 AAATTTTATTAATATATGAATTCTTATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 133

Qy     338 TGCAATAGTTTAAATAAACTTATACTTAACTTTTAAAAAAAAAAAAAAAAAAAAAAAAA 397
             || |    ||| |||   | |   ||     ||||||||||||||||||||||||
Db     132 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 73

Qy     398 AAAAAAAAA 406
           |||||||||
Db      72 AAAAAAAAA 64


RESULT 14
US-09-960-352-5187
; Sequence 5187, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
;  APPLICANT: Warren, Wesley C.
;  APPLICANT:  Tao, Nengbing
;  APPLICANT:  Byatt, John C.
;  APPLICANT:  Mathialagan, Nagappan
;  TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION ANI
;  TITLE OF INVENTION:  MUSCLE AND FAT DEPOSITION
;  FILE REFERENCE: 16511.006/37-21(10298)C
;  CURRENT APPLICATION NUMBER: US/09/960,352
;  CURRENT FILING DATE:  2001-09-24
;  NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 5187
;   LENGTH: 393
;   TYPE: DNA
;   ORGANISM: Bos taurus
;   OTHER INFORMATION: Clone ID: 22-LIB34-074-Q1-E1-F5
US-09-960-352-5187

  Query Match            13.0%;  Score 52.8;  DB 10;  Length 393;
  Best Local Similarity  63.3%;  Pred. No. 0.052;
  Matches   81;  Conservative    0;  Mismatches   47;  Indels    0;  Gaps    0;

Qy     279 CAGTTTCCAATTAAAGTCATTTCGAGCCATACTGAATAATTTTGTAATCTAACAACAGAT 338
           ||| ||||||| |||  ||  | |      | ||||| ||       ||   || || | |
Db     259 CAGGTTCCAATAAAAACCAAATTGTAGACTCCTGAAAAAAAAAAAAAAAAAAAAAAAAAA 318

Qy     339 GCAATAGTTTAAATAAACTTATACTTAACTTTTAAAAAAAAAAAAAAAAAAAAAAAAAA 398
             || |    ||| |||   | |   ||     ||||||||||||||||||||||||
Db     319 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 378

Qy     399 AAAAAAAA 406
           ||||||||
Db     379 AAAAAAAA 386
```

```
RESULT 15
US-09-918-995-13971
; Sequence 13971, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
;  APPLICANT: Hyseq, Inc.
;  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
;  TITLE OF INVENTION:  FROM VARIOUS cDNA LIBRARIES
;  FILE REFERENCE: 20411-756
;  CURRENT APPLICATION NUMBER: US/09/918,995
;  CURRENT FILING DATE:  2001-07-30
;  PRIOR APPLICATION NUMBER: US/09/235,076
;  PRIOR FILING DATE: 1999-01-20
;  NUMBER OF SEQ ID NOS: 38054
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 13971
;   LENGTH: 454
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(454)
;   OTHER INFORMATION: n = A,T,C or G
US-09-918-995-13971

  Query Match             13.0%;  Score 52.8;  DB 9;  Length 454;
  Best Local Similarity   60.4%;  Pred. No. 0.055;
  Matches   87;  Conservative    0;  Mismatches   57;  Indels    0;  Gaps    0;

Qy     263 TTTTATTGAATATCGTCAGTTTCCAATTAAAGTCATTTCGAGCCATACTGAATAATTTTG 322
           |||||   | | || |        ||||  |    | |  ||   |||    | |||||
Db     244 TTTTAACCATTTTCATGGTGGAAGAATTTTATATTTATGCAGTTGTACAATTTTATTTTT 303

Qy     323 TAATCTAACAACAGATGCAATAGTTTAAATAAACTTATACTTAACTTTTAAAAAAAAAAA 382
           |  |  || || |  || |||   | ||||||||  |          ||| ||||||||||
Db     304 TTCTGCAAGAAAAAGTGTAATGTATGAAATAAACCAAAGTCACTTGTTTGAAAAAAAAAA 363

Qy     383 AAAAAAAAAAAAAAAAAAAAAAAA 406
           ||||||||||||||||||||||||
Db     364 AAAAAAAAAAAAAAAAAAAAAAAA 387


Search completed: May  5, 2003, 21:11:13
Job time : 145 secs
```

GenCore version 5.1.5

OM nucleic - nucleic search, using sw model

Run on: May 5, 2003, 18:05:58 ; Search time 1569 Seconds
(without alignments)
4190.802 Million cell updates/sec

Title: US-09-829-481-3
Perfect score: 406
Sequence: 1 ctctactacaatcactaagt.....aaaaaaaaaaaaaaaaaaaa 406

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	62.4	15.4	355	13	BI323340	BI323340 kt69h08.y
	2	62.2	15.3	328	13	BM154538	BM154538 fv86c10.y
c	3	62.2	15.3	329	9	AL513719	AL513719 AL513719
	4	61.8	15.2	376	13	BM522187	BM522187 ESSU0725
	5	61.2	15.1	759	17	CNS06QXV	AL411257 T7 end of
	6	60.8	15.0	181	9	AI529496	AI529496 va52d02.y
	7	60.2	14.8	470	13	BG928759	BG928759 HNC56-1-C
	8	60.2	14.8	605	13	BG926805	BG926805 HNC24-1-B
	9	60.2	14.8	1101	17	CNS0021J	AL061936 Drosophil
	10	59.8	14.7	284	13	BI745513	BI745513 rk81f07.y
c	11	59.8	14.7	503	9	AL513809	AL513809 AL513809
	12	59.2	14.6	186	14	C84834	C84834 C84834 Dict
	13	59	14.5	326	13	BM154844	BM154844 fv90c01.y
	14	59	14.5	329	9	AL513719	AL513719 AL513719
	15	59	14.5	424	13	BM263207	BM263207 if34f12.y
	16	58.8	14.5	455	13	BI073560	BI073560 kt32c07.y
	17	58.2	14.3	265	14	C91235	C91235 C91235 Dict
	18	58.2	14.3	631	10	AV729934	AV729934 AV729934
	19	58	14.3	1182	14	BM906534	BM906534 AGENCOURT
	20	57.8	14.2	427	14	BQ638293	BQ638293 hd20g09.y
	21	57.8	14.2	787	17	CNS049YD	AL281182 Tetraodon
	22	57.8	14.2	1043	17	CNS0145P	AL103735 Drosophil
c	23	57.8	14.2	1101	17	CNS017V2	AL108536 Drosophil
	24	57.6	14.2	256	14	BM966539	BM966539 ko08g09.y
	25	57.4	14.1	348	13	BM186008	BM186008 fv96c02.y
	26	57.4	14.1	366	10	AW101676	AW101676 sd68f05.y
	27	57.4	14.1	425	9	AL514791	AL514791 AL514791
	28	57.4	14.1	943	12	BG392408	BG392408 602410578
	29	57.2	14.1	191	14	BQ637334	BQ637334 he08f12.y
	30	57.2	14.1	453	10	BE578190	BE578190 rk09a02.y
	31	57.2	14.1	866	13	BM015504	BM015504 603641716
	32	57.2	14.1	887	12	BG166263	BG166263 602345444
	33	57.2	14.1	924	17	CNS07A5L	AL436159 T7 end of
	34	57	14.0	438	9	AL634908	AL634908 AL634908
	35	56.8	14.0	205	10	BE666934	BE666934 150864 MA
	36	56.8	14.0	481	9	AU037107	AU037107 AU037107
	37	56.6	13.9	633	10	AW645696	AW645696 cm55h03.w
	38	56.4	13.9	166	14	BQ421843	BQ421843 AGENCOURT
	39	56.4	13.9	209	14	BM882890	BM882890 rb30g09.y
	40	56.4	13.9	298	13	BI703836	BI703836 kx32a05.y
	41	56.4	13.9	332	13	BM154632	BM154632 fv87e01.y
	42	56.4	13.9	372	13	BM089774	BM089774 503615 MA
c	43	56.4	13.9	626	9	AL513937	AL513937 AL513937
c	44	56.4	13.9	784	9	AL048397	AL048397 DKF2p586G
	45	56.2	13.8	349	9	AU033553	AU033553 AU033553

ALIGNMENTS

RESULT 1
BI323340
LOCUS BI323340 355 bp mRNA linear EST 30-JUL-2001
DEFINITION kt69h08.y1 Strongyloides ratti L2 pAMP1 v1 Chiapelli McCarter Strongyloides ratti cDNA 5' similar to TR:Q18611 Q18611 CODED FOR BY C. ELEGANS CDNA CEESK63F. [1] ;contains element MER30 repetitive element ;, mRNA sequence.
ACCESSION BI323340
VERSION BI323340.1 GI:15002526
KEYWORDS EST.
SOURCE Strongyloides ratti.
ORGANISM Strongyloides ratti
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Panagrolaimoidea; Strongyloididae; Strongyloides.
REFERENCE 1 (bases 1 to 355)
AUTHORS McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J., Wylie,T., Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Franklin,C., Tsagareishvili,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C., Underwood,K., Steptoe,M., Allen,M., Person,B., Swaller,T., Harvey,N., Schurk,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
TITLE The Washington Univ. Nematode EST Project, 1999
JOURNAL Unpublished (1999)
COMMENT Contact: McCarter JP
The Washington Univ. Nematode EST Project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800

Fax: 314 286 1810
Email: est@watson.wustl.edu
The library was constructed by Brandi Chiapelli and Dr. James McCarter (bchiapel@watson.wustl.edu & jmccarte@watson.wustl.edu) at Washington University, St. Louis. DNA Sequencing by: Washington University Genome Sequencing Center St. Louis.
Seq primer: -40RP from Gibco
High quality sequence stop: 347.

FEATURES Location/Qualifiers
source 1. .355
/organism="Strongyloides ratti"
/db_xref="taxon:34506"
/clone_lib="Strongyloides ratti L2 pAMP1 v1 Chiapelli McCarter"
/dev_stage="L2"
/lab_host="DH10B"
/note="Vector: pAMP1 (Gibco); The library was constructed by Brandi Chiapelli and Dr. James McCarter at Washington University, St. Louis. The cDNA was made by using Dynabead oligo-dT priming (Dynal). PCR based library using a modified protocol from the SMART PCR cDNA Synthesis Kit from Clontech. Directionally cloned into the UDG sites of pAMP1. Nematodes were provided by Dr. Mark Viney of Bristol, UK."

BASE COUNT 194 a 28 c 29 g 104 t
ORIGIN

Query Match 15.4%; Score 62.4; DB 13; Length 355;
Best Local Similarity 57.0%; Pred. No. 7.8;
Matches 114; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

```
Qy    207 CAACCTGCGTTTGCTATAGGTGAAAATCCGATTTATTTGCCATAATGGAGACCCGTTTTT 266
          |||    |||||    ||    |||     | | ||    || ||   |      |||||
Db    142 CAAGTATCGTTTTGGTTAACAAAAATATATACTAATAAAACAAAAAAAAATTAATTTTTT 201

Qy    267 ATTGAATATCGTCAGTTTCCAATTAAAGTCATTTCGAGCCATACTGAATAATTTTGTAAT 326
           || ||||   |   ||  || ||||    ||||  | | | |   | ||||| || || |
Db    202 TTTAAATAAATTTTATTAACACTTAATAATATTTTTATCAAAATAAATTAATATTATATT 261

Qy    327 CTAACAACAGATGCAATAGTTTAAATAAACTTATACTTAACTTTTAAAAAAAAAAAAAAA 386
           || || | |   || |    ||| |||   | |   ||     ||||||||||||||||
Db    262 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 321

Qy    387 AAAAAAAAAAAAAAAAAAA 406
          |||||||||||||||||||
Db    322 AAAAAAAAAAAAAAAAAAA 341
```

RESULT 2
BM154538
LOCUS BM154538 328 bp mRNA linear EST 03-DEC-2001
DEFINITION fv86c10.y1 Zebrafish SJD adult male II Danio rerio cDNA clone 5543827 5', mRNA sequence.
ACCESSION BM154538
VERSION BM154538.1 GI:17236719
KEYWORDS EST.
SOURCE zebrafish.
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes ; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 328)
AUTHORS Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy ,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood ,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
TITLE WashU Zebrafish EST Project 1998
JOURNAL Unpublished (1998)
COMMENT Contact: Stephen L. Johnson
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: zbrafish@watson.wustl.edu
cDNA Library construction by: Joe Barnes and Steve Johnson. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: Research Genetics web address: http://www.researchgenetics.com/
Seq primer: T3 ET from Amersham
High quality sequence stop: 322.

FEATURES Location/Qualifiers
source 1. .328
/organism="Danio rerio"
/strain="SJD"
/db_xref="taxon:7955"
/clone="5543827"
/clone_lib="Zebrafish SJD adult male II"
/sex="male"
/tissue_type="whole body"
/lab_host="DH10B"
/note="Vector: pAMP1; Site_1: EcoRI; Site_2: NotI; First strand cDNA synthesis was primed using oligo-dT on magnetic beads with an additional primer 5'-gcggccgctaatacgactcacta-taggg-3'. Second strand synthesis was a 3-cycle PCR using the primers 5'-ggccgctaatacgactcactatag-3' and 5'-aagcagtggtaacaacgcagagtacttt-ttttttttttttvn-3'. The cDNA was subsequently amplified in a 7-cycle PCR with the following primers: 5'-ggccgctaatacgactcactatag-3' and 5'-aagcagtggt-aacaacgcag. Deoxy-UMP adaptors were added in a third PCR (5 cycles) and the primers 5'-caucaucaucauggccgctaatacgactcactataggg-3' and 5'-cuacuacuacuaaagcagtggtaacaacgcagagtac-3'. Ends were treated with uracil DNA glycosylase and product with 3' overhangs was annealed to complementary ends of pAMP1. Insert can be excised using EcoRI and NotI. Library constructed by Joe Barnes and Steve Johnson (Washington University)."

BASE COUNT 188 a 32 c 39 g 69 t
ORIGIN

Query Match 15.3%; Score 62.2; DB 13; Length 328;
Best Local Similarity 53.5%; Pred. No. 8.7;
Matches 130; Conservative 0; Mismatches 113; Indels 0; Gaps 0;

```
Qy    164 AGTATTCGTCGTAGAGGAGGCTTTTGCAGAGGAACTTTCAGGACAACCTGCGTTTGCTAT 223
          ||||||  |  || | ||  |       |       |   | |   ||  ||   || | ||
Db     75 AGTATTAATTATAAAAGACACCCGGAGACTACTGCACACTGTCTAATGTGTAATTTCCAT 134

Qy    224 AGGTGAAAATCCGATTTATTTGCCATAATGGAGACCCGTTTTTATTGAATATCGTCAGTT 283
           |||   ||||   |    |   || || | |||||   | |       |  |||| |
Db    135 TGGTATTAATCATTTACTGTGTACAAAAGGCAGACCACCTGTCTAATGGAAAAGTCAATA 194

Qy    284 TCCAATTAAAGTCATTTCGAGCCATACTGAATAATTTTGTAATCTAACAACAGATGCAAT 343
            ||   |||| ||||  ||   | |   || || || | ||   || || | |   ||
Db    195 AACATGAAAAGACATTATGATGGAAAAAAAAAAAATTAGAAAAAAAAAAAAAAAAAAAAA 254

Qy    344 AGTTTAAATAAACTTATACTTAACTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 403
          |    ||| |||   | |   ||     ||||||||||||||||||||||||||||||||
Db    255 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 314

Qy    404 AAA 406
          |||
Db    315 AAA 317
```

RESULT 3
AL513719/c
LOCUS AL513719 329 bp mRNA linear EST 13-FEB-2001
DEFINITION AL513719 LTI_NFL006_PL2 Homo sapiens cDNA clone CL0BA007ZH01 3 prime, mRNA sequence.
ACCESSION AL513719
VERSION AL513719.1 GI:12777213
KEYWORDS EST.

```
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 329)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1. .329
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CL0BA007ZH01"
                     /clone_lib="LTI_NFL006_PL2"
                     /tissue_type="placenta"
                     /note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer.  Five prime end
                     enriched, double-stranded cDNA was digested with Not I and
                     cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                     vector. Library was normalized. Library was constructed by
                     Life Technologies. Contact : Feng Liang Life Technologies,
                     a division of Invitrogen 9800 Medical Center Drive
                     Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
                     Email : fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT      186 a     13 c      3 g      60 t      67 others
ORIGIN
```

```
  Query Match           15.3%;  Score 62.2;  DB 9;  Length 329;
  Best Local Similarity   44.1%;  Pred. No. 8.7;
  Matches   64;  Conservative   39;  Mismatches   42;  Indels    0;  Gaps    0;

Qy     262 TTTTTATTGAATATCGTCAGTTTCCAATTAAAGTCATTTCGAGCCATACTGAATAATTTT 321
           |||||  ||   | | :|   |||    ||    |  |||  :    :: : : |: :|:|
Db     171 TTTTTTTTTTTTTTTKTTTTTTTTTTTTTTTTTTTTTTTTTTWTWWTWWTWTWTTWTWTWT 112

Qy     322 GTAATCTAACAACAGATGCAATAGTTTAAATAAACTTATACTTAACTTTTAAAAAAAAAA 381
           |::| |:: :| : ::  :|:: ::|::::::: ::::| ::||  :: ||||||||||
Db     111 WTWWTTTWWAWAWWWWWWWWAWWWWWTWWWWWWWWWWAWWWAAAAWWAAAAAAAAAA 52

Qy     382 AAAAAAAAAAAAAAAAAAAAAAAA 406
           ||||:|:||||||||||||||||||
Db      51 AAAAWAWAAAAAAAAAAAAAAAAA 27
```

```
RESULT 4
BM522187
LOCUS       BM522187                376 bp    mRNA    linear   EST 19-FEB-2002
DEFINITION  ESSU0725 S.scabiei cDNA library Sarcoptes scabiei cDNA clone
            SAS0859 5', mRNA sequence.
ACCESSION   BM522187
VERSION     BM522187.1  GI:18706427
KEYWORDS    EST.
SOURCE      Sarcoptes scabiei.
  ORGANISM  Sarcoptes scabiei
            Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
            Acariformes; Sarcoptiformes; Astigmata; Sarcoptoidea; Sarcoptidae;
            Sarcoptes.
REFERENCE   1  (bases 1 to 376)
  AUTHORS   Ljunggren,E.L., Nilsson,D., Naslund,K. and Mattsson,J.G.
  TITLE     Expressed sequence tag analysis of the parasitic mite Sarcoptes
            scabiei
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Mattsson J.G.
            Department of Parasitology (SWEPAR)
            National Veterinary Institute
            SE-751 89 Uppsala, Sweden
            Tel: +46 18 674120
            Fax: +46 18
            Email: jens.mattsson@sva.se
            Similar to gi|8777580|dbj|BAA97098.1| (AP002460) gene_id:F1D9.26
            unknown protein [Arabidopsis thaliana]
            Seq primer: T3 primer
            High quality sequence stop: 376.
FEATURES             Location/Qualifiers
     source          1. .376
                     /organism="Sarcoptes scabiei"
                     /db_xref="taxon:52283"
                     /clone="SAS0859"
                     /clone_lib="S.scabiei cDNA library"
                     /note="The Sarcoptes scabiei mixed lifestage library was
                     constructed by Jens G Mattsson. cDNAs were synthesized
                     from poly(A)+ RNA by oligo d(T) priming, size-selected and
                     directionally cloned into the Uni-ZAP lambda vector
                     (Stratagene). The primary library was amplified on
                     XL1-Blue MRF' cells."
BASE COUNT      165 a     48 c     60 g     97 t      6 others
ORIGIN
```

```
  Query Match           15.2%;  Score 61.8;  DB 13;  Length 376;
  Best Local Similarity   57.5%;  Pred. No. 9;
  Matches  111;  Conservative    0;  Mismatches   82;  Indels    0;  Gaps    0;

Qy     214 CGTTTGCTATAGGTGAAAATCCGATTTATTTGCCATAATGGAGACCCGTTTTTATTGAAT 273
           || |||    |    ||| || |||| | ||   |   |  |||| |   || | ||  |
Db     141 CGATTGTGGAATTATAAATTCTGATTAAATTTAAACTCTAAAGACACTGATTCACTGGGT 200

Qy     274 ATCGTCAGTTTCCAATTAAAGTCATTTCGAGCCATACTGAATAATTTTGTAATCTAACAA 333
           |  | | |||    || || ||   |  |    |   |||    || | |||  || ||
Db     201 TTTTTGATTTTGTTTTTCAAATCGGGTAAAAATTTCTCGAAACTATTCGAAATAAAAAAA 260

Qy     334 CAGATGCAATAGTTTAAATAAACTTATACTTAACTTTTAAAAAAAAAAAAAAAAAAAAAA 393
           |  || || ||||| |||   | |   ||     ||||||||||||||||||||||
Db     261 GTATTTAAAAAGATTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 320

Qy     394 AAAAAAAAAAAAA 406
           |||||||||||||
Db     321 AAAAAAAAAAAAA 333
```

```
RESULT 5
CNS06QXV
LOCUS       CNS06QXV                759 bp    DNA     linear   GSS 05-JUL-2001
DEFINITION  T7 end of clone AW0AA009H09 of library AW0AA from strain CLIB 89 of
            Yarrowia lipolytica, genomic survey sequence.
ACCESSION   AL411257
VERSION     AL411257.1  GI:12180512
KEYWORDS    GSS.
SOURCE      Yarrowia lipolytica.
  ORGANISM  Yarrowia lipolytica
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Dipodascaceae; Yarrowia.
REFERENCE   1  (bases 1 to 759)
  AUTHORS   Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
            Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
            de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
            Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
            Saurin,W., Tekaia,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
            Wincker,P. and Weissenbach,J.
  TITLE     Genomic exploration of the hemiascomycetous yeasts: 1. A set of
            yeast species for molecular evolution studies
  JOURNAL   FEBS Lett. 487 (1), 3-12 (2000)
  MEDLINE   20584711
   PUBMED   11152876
REFERENCE   2  (bases 1 to 759)
  AUTHORS   Casaregola,S., Neuveglise,C., Lepingle,A., Bon,E., Feynerol,C.,
            Artiguenave,F., Wincker,P. and Gaillardin,C.
  TITLE     Genomic exploration of the hemiascomycetous yeasts: 17. Yarrowia
            lipolytica
  JOURNAL   FEBS Lett. 487 (1), 95-100 (2000)
  MEDLINE   20584727
   PUBMED   11152892
```

```
REFERENCE   3  (bases 1 to 759)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-SEP-2000) Genoscope - Centre National de Sequencage,
            2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
            seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT     This GSS is part of a random genomic sequencing program of thirteen
            yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
            exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii,
            Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
            lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
            angusta, Debaryomyces hansenii var.  hansenii, Pichia sorbitophila,
            Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
            5 kb were prepared and both extremities were sequenced. See
            keywords for description of this sequence and for the sequence of
            the other extremity of this insert.
FEATURES             Location/Qualifiers
     source          1. .759
                     /organism="Yarrowia lipolytica"
                     /strain="CLIB 89"
                     /db_xref="taxon:4952"
                     /clone="AW0AA009H09"
                     /clone_lib="AW0AA"
                     /note="end : T7"
BASE COUNT      356 a     53 c     51 g     206 t      93 others
ORIGIN

  Query Match             15.1%;  Score 61.2;  DB 17;  Length 759;
  Best Local Similarity   49.4%;  Pred. No. 6.8;
  Matches   84;  Conservative   24;  Mismatches   62;  Indels    0;  Gaps    0;

Qy     237 ATTTATTTGCCATAATGGAGACCCGTTTTTATTGAATATCGTCAGTTTCCAATTAAAGTC 296
           ||: || |    ::||  : |  : : |::|:: |: ::  | : : || | :||||
Db     155 ATWMATATWWATWMATATWTATWMAWATWWAWWMAWAWWWATWMAWATCTATWTAAATAA 214

Qy     297 ATTTCGAGCCATACTGAATAATTTTGTAATCTAACAACAGATGCAATAGTTTAAATAAAC 356
           ||:|  |   |:|   || :||     ||   |:::| | |   || |    ||| |||
Db     215 ATWTATAWWWAWAAAAAAAWATAAAAAAAAAMAMMMAAAAAAMAAAAAAAAAAAAAAAA 274

Qy     357 TTATACTTAACTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 406
             | ||  ||   |||||||||||||:|||||||||||||||||||||:|
Db     275 AAAAACAAAAAAATTAAAAAAAAAAAAWAAAAAAAAAAAAAAAAAAAAAAWA 324


RESULT 6
AI529496
LOCUS       AI529496                181 bp    mRNA    linear   EST 18-MAR-1999
DEFINITION  va52d02.y1 Soares mouse 3NME12 5 Mus musculus cDNA clone
            IMAGE:734979 5', mRNA sequence.
ACCESSION   AI529496
VERSION     AI529496.1  GI:4443631
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 181)
  AUTHORS   Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
            Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person
            ,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter
            ,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterston,R. and Wilson,R.
  TITLE     The WashU-NCI Mouse EST Project 1999
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Marra M/WashU-NCI Mouse EST Project 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:452027
            Seq primer: -40RP from Gibco
            High quality sequence stop: 91.
FEATURES             Location/Qualifiers
     source          1. .181
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:734979"
                     /clone_lib="Soares mouse 3NME12 5"
                     /sex="unknown"
                     /tissue_type="fetus"
                     /dev_stage="12.5dpc total fetus"
                     /lab_host="DH10B"
                     /note="Organ: whole fetus; Vector: pT7T3D-Pac (Pharmacia)
                     with a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Not I - oligo(dT) primer
                     [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCCTTATTTTTTTTTTTTTTTTT
                     3'], on total mouse RNA [provided by Minoru Ko, Wayne
                     State Univ.]; double-stranded cDNA was ligated to Eco RI
                     adaptors (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of the modified pT7T3 vector.
                     Library went through one round of normalization, and was
                     constructed by Bento Soares and M. Fatima Bonaldo.  "
BASE COUNT      110 a     19 c     18 g     34 t
ORIGIN

  Query Match             15.0%;  Score 60.8;  DB 9;  Length 181;
  Best Local Similarity   80.7%;  Pred. No. 20;
  Matches   71;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy     319 TTTGTAATCTAACAACAGATGCAATAGTTTAAATAAACTTATACTTAACTTTTAAAAAAA 378
           |||| | | | |||| ||||| |||| |||  ||||    || ||||||  | |||||||
Db      54 TTTGAACTTTCACAAGAGATGTAATAATTTTGATAATAAAATTCTTAACCATAAAAAAAA 113

Qy     379 AAAAAAAAAAAAAAAAAAAAAAAAAAAAA 406
           |||||||||||||||||||||||||||||
Db     114 AAAAAAAAAAAAAAAAAAAAAAAAAAAAA 141


RESULT 7
BG928759
LOCUS       BG928759                470 bp    mRNA    linear   EST 06-NOV-2001
DEFINITION  HNC56-1-C12.R HNC (Human Normal Cartilage) Homo sapiens cDNA, mRNA
            sequence.
ACCESSION   BG928759
VERSION     BG928759.1  GI:14323282
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 470)
  AUTHORS   Kumar,S., Connor,J.R., Dodds,R.A., Halsey,W., Van Horn,M., Mao,J.,
            Sathe,G., Mui,P., Agarwal,P., Badger,A.M., Lee,J.C., Gowen,M. and
            Lark,M.W.
  TITLE     Identification and initial characterization of 5000 expressed
            sequenced tags (ESTs) each from adult human normal and
            osteoarthritic cartilage cDNA libraries
  JOURNAL   Osteoarthr. Cartil. 9 (7), 641-653 (2001)
  MEDLINE   21482651
COMMENT     Contact: Sanjay Kumar
            UW2109
            GlaxoSmithKline
            709 Swedeland Road, P.O. Box 1539, King of Prussia, PA 19406, USA
            Tel: 610-270-7245
            Fax: 610-270-5598
            Email: sanjay_kumar-1@gsk.com
            Seq primer: T7.
FEATURES             Location/Qualifiers
     source          1. .470
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="HNC (Human Normal Cartilage)"
```

```
                    /tissue_type="cartilage"
                    /lab_host="E.coli DH10 B"
                    /note="Vector: pSPORT I; Site_1: SalI; Site_2: NotI;
                    Directional"
BASE COUNT      230 a      82 c      67 g      91 t
ORIGIN

  Query Match             14.8%;  Score 60.2;  DB 13;  Length 470;
  Best Local Similarity   63.4%;  Pred. No. 13;
  Matches   92;  Conservative    0;  Mismatches   53;  Indels    0;  Gaps    0;

Qy     262 TTTTTATTGAATATCGTCAGTTTCCAATTAAAGTCATTTCGAGCCATACTGAATAATTTT 321
           || ||||  |  |     | |||||| || |||| ||| |   | |||   |  ||||||
Db     238 TTATTATGCATAAAGCCAAATTTCCAGTTTAAGTAATTGCCTACAATAAAAAGAAATTTT 297

Qy     322 GTAATCTAACAACAGATGCAATAGTTTAAATAAACTTATACTTAACTTTTAAAAAAAAAA 381
           |    | || || | |   || |    ||| |||   | |   ||     ||||||||||
Db     298 GCCTGCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 357

Qy     382 AAAAAAAAAAAAAAAAAAAAAAAA 406
           ||||||||||||||||||||||||
Db     358 AAAAAAAAAAAAAAAAAAAAAAAA 382


RESULT 8
BG926805
LOCUS       BG926805                605 bp    mRNA    linear   EST 06-NOV-2001
DEFINITION  HNC24-1-B1.R HNC (Human Normal Cartilage) Homo sapiens cDNA, mRNA
            sequence.
ACCESSION   BG926805
VERSION     BG926805.1  GI:14321328
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 605)
  AUTHORS   Kumar,S., Connor,J.R., Dodds,R.A., Halsey,W., Van Horn,M., Mao,J.,
            Sathe,G., Mui,P., Agarwal,P., Badger,A.M., Lee,J.C., Gowen,M. and
            Lark,M.W.
  TITLE     Identification and initial characterization of 5000 expressed
            sequenced tags (ESTs) each from adult human normal and
            osteoarthritic cartilage cDNA libraries
  JOURNAL   Osteoarthr. Cartil. 9 (7), 641-653 (2001)
  MEDLINE   21482651
COMMENT     Contact: Sanjay Kumar
            UW2109
            GlaxoSmithKline
            709 Swedeland Road, P.O. Box 1539, King of Prussia, PA 19406, USA
            Tel: 610-270-7245
            Fax: 610-270-5598
            Email: sanjay_kumar-1@gsk.com
            Seq primer: T7.
FEATURES             Location/Qualifiers
     source          1. .605
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="HNC (Human Normal Cartilage)"
                     /tissue_type="cartilage"
                     /lab_host="E.coli DH10 B"
                     /note="Vector: pSPORT I; Site_1: SalI; Site_2: NotI;
                     Directional"
BASE COUNT      261 a     122 c      99 g     123 t
ORIGIN

  Query Match             14.8%;  Score 60.2;  DB 13;  Length 605;
  Best Local Similarity   63.4%;  Pred. No. 11;
  Matches   92;  Conservative    0;  Mismatches   53;  Indels    0;  Gaps    0;

Qy     262 TTTTTATTGAATATCGTCAGTTTCCAATTAAAGTCATTTCGAGCCATACTGAATAATTTT 321
           || ||||  |  |     | |||||| || |||| ||| |   | |||   |  ||||||
Db     205 TTATTATGCATAAAGCCAAATTTCCAGTTTAAGTAATTGCCTACAATAAAAAGAAATTTT 264

Qy     322 GTAATCTAACAACAGATGCAATAGTTTAAATAAACTTATACTTAACTTTTAAAAAAAAAA 381
           |    | || || | |   || |    ||| |||   | |   ||     ||||||||||
Db     265 GCCTGCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 324

Qy     382 AAAAAAAAAAAAAAAAAAAAAAAA 406
           ||||||||||||||||||||||||
Db     325 AAAAAAAAAAAAAAAAAAAAAAAA 349


RESULT 9
CNS0021J
LOCUS       CNS0021J                1101 bp    DNA     linear   GSS 03-JUN-1999
DEFINITION  Drosophila melanogaster genome survey sequence TET3 end of BAC #
            BACR05N11 of RPCI-98 library from Drosophila melanogaster (fruit
            fly), genomic survey sequence.
ACCESSION   AL061936
VERSION     AL061936.1  GI:4940214
KEYWORDS    GSS.
SOURCE      Drosophila melanogaster.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 1101)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     Determination of this BAC-end sequence was carried out as part of a
            collaboration with the Berkeley Drosophila Genome Project (BDGP).
            The BDGP is constructing a physical map of the Drosophila
            melanogaster genome using these BACs. For further information
            please see  http://www.fruitfly.org The BDGP Drosophila
            melanogaster BAC library was prepared by Kazutoyo Osoegawa and
            Aaron Mammoser in Pieter de Jong's laboratory in the Department of
            Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
            NY. The library is named RPCI-98 and was constructed by partial
            EcoRI digestion of Drosophila DNA provided by the BDGP from the
            isogenic strain y2; cn bw sp, the same strain used for the BDGP's
            P1 and EST libraries. A more detailed description of the library
            and how to order individual BAC clones, the entire library, or
            filters for hybridization from the BACPAC Resource Center can be
            found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES             Location/Qualifiers
     source          1. .1101
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="BACR05N11"
                     /clone_lib="RPCI-98"
                     /note="end : TET3"
BASE COUNT      631 a       7 c      28 g     289 t     146 others
ORIGIN

  Query Match             14.8%;  Score 60.2;  DB 17;  Length 1101;
  Best Local Similarity   52.0%;  Pred. No. 7.2;
  Matches   92;  Conservative   16;  Mismatches   69;  Indels    0;  Gaps    0;

Qy     230 AAATCCGATTTATTTGCCATAATGGAGACCCGTTTTTATTGAATATCGTCAGTTTCCAAT 289
           |:|    :||||::   | | :  |         :| ||  |: |  |   |||  ||:
Db     666 AWAAWTTWTTTATWWATWAAATWAAATTTTAAAAAWTTTTTTAWTTTTTTTWTTTAAAAW 725

Qy     290 TAAAGTCATTTCGAGCCATACTGAATAATTTTGTAATCTAACAACAGATGCAATAGTTTA 349
           |  | | :|||  |   ||: : ::||| ||| | |: ||: || |     ||   |||
Db     726 TTTATTAWTTTTAAAAWATWAWTWWTAAATTTTTTAWATAWTAAWATTAWAAAATTTTTT 785

Qy     350 AATAAACTTATACTTAACTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 406
           |||  | || |  |||| |   ||||||||||||||:||| ||||||| |||||||
Db     786 AATTTAATTTTTTTTAATTAAAAAAAAAAAAAAAAAAAWAAATAAAAAAATAAAAAAA 842


RESULT 10
BI745513
```

```
LOCUS       BI745513                284 bp    mRNA    linear   EST 25-SEP-2001
DEFINITION  rk81f07.y3 Meloidogyne javanica egg pAMP1 v6 Chiapelli McCarter
            Meloidogyne javanica cDNA 5', mRNA sequence.
ACCESSION   BI745513
VERSION     BI745513.1  GI:15767322
KEYWORDS    EST.
SOURCE      root-knot nematode.
  ORGANISM  Meloidogyne javanica
            Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
            Tylenchoidea; Heteroderidae; Meloidogyninae; Meloidogyne.
REFERENCE   1  (bases 1 to 284)
  AUTHORS   McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J., Wylie,T.
            , Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B., Bowers,Y.
            , Gibbons,M., Ritter,E., Bennett,J., Franklin,C., Tsagareishvili,R.
            , Ronko,I., Kennedy,S., Maguire,L., Beck,C., Underwood,K., Steptoe
            ,M., Allen,M., Person,B., Swaller,T., Harvey,N., Schurk,R., Kohn,S.
            , Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and
            Wilson,R.
  TITLE     The Washington Univ. Nematode EST Project, 1999
  JOURNAL   Unpublished (1999)
COMMENT     Contact: McCarter JP
            The Washington Univ. Nematode EST Project, 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            The library was constructed by Brandi Chiapelli and Dr. James
            McCarter (bchiapel@watson.wustl.edu & jmccarte@watson.wustl.edu) at
            Washington University, St. Louis. DNA Sequencing by: Washington
            University Genome Sequencing Center St. Louis.
            Putative full length read
            The vector to vector length is 285.
FEATURES             Location/Qualifiers
     source          1. .284
                     /organism="Meloidogyne javanica"
                     /db_xref="taxon:6303"
                     /clone_lib="Meloidogyne javanica egg pAMP1 v6 Chiapelli
                     McCarter"
                     /dev_stage="enriched for eggs"
                     /lab_host="DH10B"
                     /note="Vector: pAMP1 (Gibco); The library was constructed
                     by Brandi Chiapelli and Dr. James  McCarter at Washington
                     University, St. Louis. The cDNA was made by using Dynabead
                     oligo-dT priming (Dynal). PCR based library using a
                     modified  protocol from the SMART PCR cDNA Synthesis Kit
                     from Clontech. Directionally  cloned into the UDG sites of
                     pAMP1. Nematodes were provided by Dr. David  Bird of North
                     Carolina State University."
BASE COUNT      140 a     29 c     38 g     77 t
ORIGIN
```

```
  Query Match             14.7%;  Score 59.8;  DB 13;  Length 284;
  Best Local Similarity   58.9%;  Pred. No. 20;
  Matches  103;  Conservative    0;  Mismatches   72;  Indels    0;  Gaps    0;

Qy     232 ATCCGATTTATTTGCCATAATGGAGACCCGTTTTTATTGAATATCGTCAGTTTCCAATTA 291
           | || |||  | | ||  |    | | | | | || |||||  |     |   | ||||
Db     101 AACCAATTGTTCTTCCTCAGAATATAACAGATATTGTTGAAACTACAACGAAGCAAATTT 160

Qy     292 AAGTCATTTCGAGCCATACTGAATAATTTTGTAATCTAACAACAGATGCAATAGTTTAAA 351
           ||   || |  |   | | ||||||| ||   |||| | |    |  | | ||   |||||
Db     161 AATAAATATTAAATAAAATTGAATATTTAACTAATTTGAATGTACTTACTATTAATTAAA 220

Qy     352 TAAACTTATACTTAACTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 406
           ||    ||  | ||   || ||  |||||||||||||||||||||||||||||||
Db     221 AAATAAAATTTTAAAAACTTTAAGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 275
```

```
RESULT 11
AL513809/c
LOCUS       AL513809                503 bp    mRNA    linear   EST 13-FEB-2001
DEFINITION  AL513809 LTI_NFL006_PL2 Homo sapiens cDNA clone CL0BA0092F09 3
            prime, mRNA sequence.
ACCESSION   AL513809
VERSION     AL513809.1  GI:12777303
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 503)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1. .503
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CL0BA009ZF09"
                     /clone_lib="LTI_NFL006_PL2"
                     /tissue_type="placenta"
                     /note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer.  Five prime end
                     enriched, double-stranded cDNA was digested with Not I and
                     cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                     vector. Library was normalized. Library was constructed by
                     Life Technologies. Contact : Feng Liang Life Technologies,
                     a division of Invitrogen 9800 Medical Center Drive
                     Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
                     Email : fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT      138 a     144 c      4 g     133 t      84 others
ORIGIN
```

```
  Query Match             14.7%;  Score 59.8;  DB 9;  Length 503;
  Best Local Similarity   42.8%;  Pred. No. 14;
  Matches   62;  Conservative   40;  Mismatches   43;  Indels    0;  Gaps    0;

Qy     262 TTTTTATTGAATATCGTCAGTTTCCAATTAAAGTCATTTCGAGCCATACTGAATAATTTT 321
           |||||  ||:  : : :: ::|||   ||::|: : :||:  :   ::  | ::::::::::
Db     208 TTTTTTTTKGKKGKGKKKDKTTTDTTATWWAWWWWWTTWWWWWWWWWTWTTWWWWWWWWW 149

Qy     322 GTAATCTAACAACAGATGCAATAGTTTAAATAAACTTATACTTAACTTTTAAAAAAAAAA 381
           :::: ::: :: : ::  :::|    ||| |||   | |   ||     ||||||||||
Db     148 WWWWWTWWWWWWWWWWWWWWWAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 89

Qy     382 AAAAAAAAAAAAAAAAAAAAAAAA 406
           ||||||||||||||||||||||||
Db      88 AAAAAAAAAAAAAAAAAAAAAAAA 64
```

```
RESULT 12
C84834
LOCUS       C84834                  186 bp    mRNA    linear   EST 12-JUL-1999
DEFINITION  C84834 Dictyostelium discoideum SS (H.Urushihara) Dictyostelium
            discoideum cDNA clone SSG726, mRNA sequence.
ACCESSION   C84834
VERSION     C84834.1  GI:2864981
KEYWORDS    EST.
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 186)
  AUTHORS   Morio,T., Urushihara,H., Saito,T., Ugawa,Y., Mizuno,H., Yoshida,M.,
            Yoshino,R., Mitra,B.N., Pi,M., Sato,T., Takemoto,K., Yasukawa,H.,
            Williams,J., Maeda,M., Takeuchi,I., Ochiai,H. and Tanaka,Y.
  TITLE     The Dictyostelium developmental cDNA project: generation and
            analysis of expressed sequence tags from the first-finger stage of
            development
  JOURNAL   DNA Res. 5 (6), 335-340 (1998)
  MEDLINE   99156227
```

```
COMMENT     Contact: Hideko Urushihara
            Institute of Biological Sciences
            University of Tsukuba
            1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan
            Tel: 81-298-53-4664
            Fax: 81-298-53-6614
            Email: hideko@biol.tsukuba.ac.jp
            PROJECT = 'Dictyostelium discoideum cDNA project in Japan'
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .186
                     /organism="Dictyostelium discoideum"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /clone="SSG726"
                     /clone_lib="Dictyostelium discoideum SS (H.Urushihara)"
                     /dev_stage="slug"
BASE COUNT      103 a     12 c     16 g     55 t
ORIGIN

  Query Match             14.6%;  Score 59.2;  DB 14;  Length 186;
  Best Local Similarity   76.0%;  Pred. No. 33;
  Matches   73;  Conservative    0;  Mismatches   23;  Indels    0;  Gaps    0;

Qy     311 TGAATAATTTTGTAATCTAACAACAGATGCAATAGTTTAAATAAACTTATACTTAACTTT 370
           | |||  | |   | | |  |||      ||||| ||||| | || ||||  ||||||||
Db      55 TAAATGGTATCCCATTATGTCAAATTCCCCAATAATTTAATTTAATTTATTTTTAACTTT 114

Qy     371 TAAAAAAAAAAAAAAAAAAAAAÃAAAAAAAAAAAAAAA 406
           |||||||||||||||||||||||||||||||||||||
Db     115 TAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 150


RESULT 13
BM154844
LOCUS       BM154844                326 bp    mRNA    linear   EST 03-DEC-2001
DEFINITION  fv90c01.y1 Zebrafish SJD adult male II Danio rerio cDNA clone
            5544193 5', mRNA sequence.
ACCESSION   BM154844
VERSION     BM154844.1  GI:17237292
KEYWORDS    EST.
SOURCE      zebrafish.
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
            ; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 326)
  AUTHORS   Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy
            ,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood
            ,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B.,
            Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E.,
            Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.
            and Wilson,R.
  TITLE     WashU Zebrafish EST Project 1998
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Stephen L. Johnson
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: zbrafish@watson.wustl.edu
            cDNA Library construction by: Joe Barnes and Steve Johnson. DNA
            Sequencing by: Washington University Genome Sequencing Center Clone
            distribution: Research Genetics web address:
            http://www.researchgenetics.com/
            Seq primer: T3 ET from Amersham
            High quality sequence stop: 322.
FEATURES             Location/Qualifiers
     source          1. .326
                     /organism="Danio rerio"
                     /strain="SJD"
                     /db_xref="taxon:7955"
                     /clone="5544193"
                     /clone_lib="Zebrafish SJD adult male II"
                     /sex="male"
                     /tissue_type="whole body"
                     /lab_host="DH10B"
                     /note="Vector: pAMP1; Site_1: EcoRI; Site_2: NotI; First
                     strand cDNA synthesis was primed using oligo-dT on
                     magnetic beads with an additional primer .
                     5'-gcggccgctaatacgactcacta-taggg-3'. Second strand
                     synthesis was a 3-cycle PCR using the primers
                     5'-ggccgctaatacgactcactatag-3' and
                     5'-aagcagtggtaacaacgcagagtacttt-ttttttttttttvn-3'. The
                     cDNA was subsequently amplified in a 7-cycle PCR with the
                     following primers: 5'-ggccgctaatacgactcactatag-3' and
                     5'-aagcagtggt-aacaacgcag. Deoxy-UMP adaptors were added in
                     a third PCR (5 cycles) and the primers
                     5'-caucaucaucauggccgctaatacgactcactataggg-3' and
                     5'-cuacuacuacuaaagcagtggtaacaacgcagagtac-3'. Ends were
                     treated with uracil DNA glycosylase and product with 3'
                     overhangs was annealed to complementary ends of pAMP1.
                     Insert can be excised using EcoRI and NotI. Library
                     constructed by Joe Barnes and Steve Johnson (Washington
                     University)."
BASE COUNT      190 a     32 c     37 g     67 t
ORIGIN

  Query Match             14.5%;  Score 59;  DB 13;  Length 326;
  Best Local Similarity   52.7%;  Pred. No. 24;
  Matches  128;  Conservative    0;  Mismatches  115;  Indels    0;  Gaps    0;

Qy     164 AGTATTCGTCGTAGAGGAGGCTTTTGCAGAGGAACTTTCAGGACAACCTGCGTTTGCTAT 223
           ||||||  |  || | ||  |      |      |    | |   ||  ||   || | ||
Db      75 AGTATTAATTATAAAAGACACCCGGAGACTACTGCACACTGTCTAATGTGTAATTTCCAT 134

Qy     224 AGGTGAAAATCCGATTTATTTGCCATAATGGAGACCCGTTTTTATTGAATATCGTCAGTT 283
           |||   ||||   |    |   || || | |||||   | |        |  |||| |
Db     135 TGGTATTAATCATTTACTGTGTACAAAAGGCAGACCACCTGTCTAATGGAAAAGTCAATA 194

Qy     284 TCCAATTAAAGTCATTTCGAGCCATACTGAATAATTTTGTAATCTAACAACAGATGCAAT 343
           ||   |||| |||| ||   | |   || ||    | ||   || || | |   ||
Db     195 AACATGAAAAGACATTATGATGGAAAAAAAAAAAAAAAGAAAAAAAAAAAAAAAAAAAA 254

Qy     344 AGTTTAAATAAACTTATACTTAACTTTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 403
           |   ||| |||   | |   ||      ||||||||||||||||||||||||||||||
Db     255 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 314

Qy     404 AAA 406
           |||
Db     315 AAA 317


RESULT 14
AL513719
LOCUS       AL513719                329 bp    mRNA    linear   EST 13-FEB-2001
DEFINITION  AL513719 LTI_NFL006_PL2 Homo sapiens cDNA clone CL0BA007ZH01 3
            prime, mRNA sequence.
ACCESSION   AL513719
VERSION     AL513719.1  GI:12777213
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 329)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1. .329
                     /organism="Homo sapiens"
```

/db_xref="taxon:9606"
/clone="CL0BA007ZH01"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 186 a 13 c 3 g 60 t 67 others
ORIGIN

Query Match 14.5%; Score 59; DB 9; Length 329;
Best Local Similarity 42.8%; Pred. No. 24;
Matches 62; Conservative 39; Mismatches 44; Indels 0; Gaps 0;

```
Qy    262 TTTTTATTGAATATCGTCAGTTTCCAATTAAAGTCATTTCGAGCCATACTGAATAATTTT 321
          ||||| ||   | |  |   |||   :|:    |  |||      :|  | ::|::::::
Db     19 TTTTTTTTTTTTTTTTTTTTTTTTTWTWTTTTTTTTTTTTTTTWWTTTTWWWTWWWWWW 78

Qy    322 GTAATCTAACAACAGATGCAATAGTTTAAATAAACTTATACTTAACTTTTAAAAAAAAAA 381
           ::::  :: :: : ::  ::|: :: ||::|:|  :|:| : :| : ::|:|||||||||
Db     79 WWWWWWAWWWWWTWWWWWWWWTWTWWAAAWWAWAWAWAAWAWAWWAWWAAAAAAAAA 138

Qy    382 AAAAAAAAAAAAAAAAAAAAAAAA 406
          ||||||||||||||||||:|||||||
Db    139 AAAAAAAAAAAAAAAAAMAAAAAA 163
```

RESULT 15
BM263207
LOCUS BM263207 424 bp mRNA linear EST 12-MAR-2002
DEFINITION if34f12.y2 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens cDNA clone IMAGE:5678374 5', mRNA sequence.
ACCESSION BM263207
VERSION BM263207.1 GI:17926247
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 424)
AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K., Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S., Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A., Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas ,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T. , Jackson,Y. and Bowers,Y.
TITLE Endocrine Pancreas Consortium
JOURNAL Unpublished (2000)
COMMENT Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Douglas Melton DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Juliana Brown (brown@fas.harvard.edu) This sequence now available from the IMAGE consortium, for clone orders contact: info@image.llnl.gov
Seq primer: -40RP from Gibco
High quality sequence stop: 423.
FEATURES Location/Qualifiers
source 1. .424
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5678374"
/clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
/sex="Both"
/tissue_type="Islets of Langerhans"
/dev_stage="Adult"
/lab_host="DH10B"
/note="Organ: Pancreas; Vector: pSPORT1; Site_1: Not 1; Site_2: Sal 1; Starting library constructed using SuperScript Plasmid Library kit (Life Technologies). cDNA made by oligo-dT priming. Size-selected by column fractionation; average insert size 1.08 kb. Library was amplified once on solid support and plasmid DNA from library was prepared. The library DNA was normalized by method #4 from Bonaldo, Lennon, and Soares 1996 Genome Research 6:791-806; 0.5 microgram single-stranded library plasmid DNA was mixed with 5 micrograms PCR product representing library inserts and hybridized to an Ecot of 20. Single-stranded (unhybridized) plasmids were isolated by hydroxyapatite chromatography and used to make this library."

BASE COUNT 219 a 49 c 39 g 116 t 1 others
ORIGIN

Query Match 14.5%; Score 59; DB 13; Length 424;
Best Local Similarity 50.7%; Pred. No. 20;
Matches 140; Conservative 0; Mismatches 136; Indels 0; Gaps 0;

```
Qy    131 TTTAATGCAGGAAAATGCCATAGACATTGCAAAAGTATTCGTCGTAGAGGAGGCTTTTGC 190
          ||| ||||| || |||||| ||  |  | | | |||   ||  | |  |   |||
Db    125 TTTTATGCAAAAATCTGCCATTGAATTGTCCATAATATGTTTCTCAAAATATCATTTATA 184

Qy    191 AGAGGAACTTTCAGGACAACCTGCGTTTGCTATAGGTGAAAATCCGATTTATTTGCCATA 250
          |       | | |     |||   | |   ||||  | ||  ||        |   |||
Db    185 ATTTATTATGTGATTTAGACCCAAGCTAAATATAAATTAACTTCATTAACTATGTACATT 244

Qy    251 ATGGAGACCCGTTTTTATTGAATATCGTCAGTTTCCAATTAAAGTCATTTCGAGCCATAC 310
          ||   ||| |  ||  |||   |      | |  | ||  ||| | |     |  |
Db    245 ATCATGACACTGTTGAATTACTTTCACAGACTGACAAAGCAAAATTAAAAATAAACTCTG 304

Qy    311 TGAATAATTTTGTAATCTAACAACAGATGCAATAGTTTAAATAAACTTATACTTAACTTT 370
          | | | |   ||  || || | |   || |    ||| |||   | |   | |
Db    305 TAAGTGACCTANAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACCCAA 364

Qy    371 TAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 406
          |||||||||||||||||||||||||||||||||||||
Db    365 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 400
```

Search completed: May 5, 2003, 20:17:32
Job time : 1575 secs